VOLUME X MARCH, 1919 NUMBER 3

THE JOURNAL OF
EDUCATIONAL PSYCHOLOGY

IS INDIVIDUAL LEARNING CAPACITY CONSTANT FOR DIFFERENT TYPES OF MATERIAL?

W. H. PYLE

University of Missouri

Learning is now defined as bond-building. In habit-formation, for example, nervous bonds or connections are established between stimulus and response. A nervous bond, of course, is merely a path of least resistance. In ideational learning bonds are established—paths of small resistance—between the cortical processes underlying the ideas involved. The question which we raise in this paper is this: are the characteristics of the nervous system such that a person may be spoken of as a "good learner" or a "poor learner;" is it easy for one person to establish new neural bonds, involved in learning, and relatively difficult for others? Must we say that a person can easily learn mathematics, but learns some other branch, let us say history, with difficulty? Or can we say that a person is a good learner or a poor learner of whatever he undertakes to learn? In the schools, for example, is a person known as a *good* student or a *poor* student in general, or is he known as a good student in history and language, for example, and a poor student in science and mathematics? Generally, the former is the case. Students are spoken of as good or poor without reference to any particular subject of study. There are, however, numerous exceptions. Sometimes a student is known as an excellent student in science and as a poor student in language. In such cases the question arises whether the student's being good in one subject and poor in another may be due to such factors as difference in the preliminary training for the subjects, or difference in interest in the two subjects. Several investigators have computed the correlation between the standing of University students in one subject with standing in another. These correlations have in all cases been found to be high, though of course not unity.

The question which we have raised is of the highest importance to education, for if learning, whatever its nature, depends upon fundamental characteristics of the brain and nervous system, then

it will be found a comparatively simple matter to determine by experiment the learning capacity of the children in the public schools and the students in the colleges. On the other hand, if learning is a specific matter, we cannot speak of learning capacity in a general sense, but must speak of ability to learn this or that.

The first experiments of the writer in this field were performed in 1911–1912. On the basis of these experiments the statement was made in *The Outlines of Educational Psychology* (p. 202) that "it seems probable that every individual has a fairly definite coefficient of learning capacity that is fairly constant." The few experiments conducted at that time consisted in having subjects commit to memory nonsense syllables. The lists contained twenty-six syllables each. The number of repetitions necessary to learn a list at one sitting was determined. The number of repetitions for re-learning a series on succeeding days was also determined. It was found that re-learning time had a very definite ratio to learning time. The experiments showed that quick learners retained well what they had learned, as determined by their later re-learning time. While the experiment threw no light on the matter of learning different types of material, it showed great definiteness and constancy in individuals in learning the same kind of material.

For several years I have been trying to discover how the matter stands with reference to different types of material. Ability in learning four different types of material has been studied; namely, substitution, nonsense syllables, card distribution, and marble distribution. These types of learning are quite different, one from the other. Substitution tests and nonsense-syllable learning are types of what may be called intellectual learning. The associations are in large measure between ideas. The motor element is unimportant in them; at any rate, is much less important than it is in the card and marble-distributing tests.

Three types of substitution tests were used; writing digits for symbols, symbols for digits, and symbols for the letters of the alphabet. In all of the substitution tests, except alphabet-symbol substitution, an experiment consisted of five minutes work. In some cases the experiment was continued for several five-minute periods. Four different keys were used.

In the learning of nonsense syllables the method was that of complete learning of a series of ten syllables each. The syllables

were exposed by means of a Jastrow tachistoscope, each syllable being exposed for a period of one second. After the exposure of the series the subjects wrote in their proper sequence all the syllables they could recall. The exposures were continued until all the syllables could be repeated in their proper order. The syllables were in all cases written down by the subjects immediately after the exposure. In computing the correlations we have in some cases considered the total number of repetitions required for the subject to get all the syllables; in some cases we used the efficiency after a certain exposure, say the second.

The card-distributing experiment consisted in distributing 150 cards into the 30 compartments of a box. The compartments were numbered from 11 to 40. The cards were numbered in the same manner, five of each number. The numbering of the boxes was without any system, *i. e.*, the number did not run consecutively on adjacent boxes. The cards were in all cases carefully shuffled before each distribution. Efficiency was of course determined by the time required for distribution.

The marble experiment consisted in distributing by means of a complicated machine, marbles of four different colors and two different sizes into their corresponding compartments. Efficiency was determined by the number of marbles correctly placed per minute.

The card and marble experiments are alike in both being motor learning, but they are quite different in detail. In the case of the card distribution, simple associations had to be built up between a number and a place. In the case of the marbles, the matter was much more complicated. Not only was it necessary to build up several types of associations, as between the color of a marble and the place where it was to be found, the color of a marble and the place where it was to be put, but also with the movements of the feet in operating pedals for different kinds of marbles. In the operation of the machine distinctions must constantly be made among the six different kinds of marbles, and four holes on the top of the table, and the movements of both feet and both hands. The experiment, therefore, demanded a high degree of co-ordination of muscular movements, and concentrated attention.

TABLE OF CORRELATIONS

The correlations which follow were all computed by the Pearson formula. The experiments on which they were based were performed at different times and upon different groups of students

		r.	P. E.
Substitution (digit-symbol and symbol-digit) with:			
Cards		.64	.074
"		.66	.076
"		.559	.066
"		.59–	.09
"		.56	.06
"	Av.....	.602	
Marbles		.469	.103
"		.54	.10
"		.62–	.058
"		.68	.036
"		.82	.072
"	Av.....	.626	
Nonsense syl.		.432	.077
Alphabet-symbol		.57	.091
" "		.59	.08
" "	Av.....	.58	
Average of substitution correlations.....		.594	
Alphabet-symbol substitution test with:			
Cards		.54	.095
"		.49	.11
"	Av.....	.515	
Substitution (digit symbol)		.57	.091
"		.59	.08
"	Av.....	.58	
Average of alphabet-symbol correlations.....		.547	
Nonsense syllables, with:			
Cards		.399	.08
Substitution		.432	.077
Mirror-writing		.505	.11
Marbles		.342	.084
"		.53	.196
	Av.....	.441	
Average all nonsense syllable correlations.....		.441	

		r.	P. E.
Card-distributing with:			
Marbles		.31	.13
"		.218	.091
"	Av.....	.264	
Substitution		.559	.066
"		.56	.06
"		.64	.074
"		.66	.076
"		.59	.09
"	Av.....	.602	
Alphabet-symbol		.54	.095
" "		.49	.11
" "	Av....	.515	
Nonsense syllables		.399	.080
Average all card correlations.....		.496	
Marble-distributing with:			
Substitution		.620	.058
"		.469	.103
"		.54	.10
"		.82	.072
"		.68	.036
"	Av.....	.626	
Nonsense syllables		.342	.084
" "		.53	.196
" "	Av.....	.436	
Cards		.31	.13
"		.218	.091
"	Av.....	.264	
Average all marble correlations.....		.503	

Average inter-correlation of all the learning tests.. .516

The ranking of the tests, based on their degree of correlation with the other tests, is as follows: digit substitution test, .594; alphabet substitution test, .547; marble distributing, .503; card distributing, .496; nonsense syllable learning, .441.

In the accompanying table the correlations are shown between each type of learning and the other types. These correlations are based upon experiments with several hundred University students, and were performed in the years 1914–1918. The groups studied vary all the way from nine to 120. The results from all of them are shown. It will be seen that the average of the "raw" correlations between one type of learning and the other types is approximately .50. The true correlation would be considerably higher. This study therefore shows a high, but not a perfect, correlation between ability for one type of learning and other types.

Two questions now arise. First, if all extraneous factors were eliminated, would the correlation be unity, showing that we can speak of learning capacity as a definite characteristic of the nervous system? Second, what is the cause of the positive correlation which we have actually found? These questions are really one, however, for if the true correlation between these types of learning is unity, there could be no question that the explanation would be a definite and constant learning capacity in the subjects.

There is considerable evidence that if all the disturbing factors could be removed, the true correlation would be much nearer unity than the raw correlation which we have actually found. The subjects were all mature men and women whose ages would average over twenty-one. Their experiences in life had been widely different; different prejudices and attitudes had been acquired. They therefore entered these experiments not only with different learning capacities, but with all sorts of differences due, not to learning capacity, but to past experience. In the marble-distributing experiment, for example, it was found that students who had had considerable experience in running typewriters, or in playing pianos, made better scores. In the card-distributing experiment not only was ability to learn the location of the boxes a factor, but facility in handling the cards was also a factor. In playing cards, some of the subjects had acquired great dexterity in handling them; others who had not played cards made poorer records because of lack of dexterity. And with each type of learning the subjects came to the experiment with differences due, not to native learning capacity, but to difference in past experience. Now, if taking the subjects just as we found them, without eliminating any of the disturbing factors, we found an inter-correlation of approximately .50, it is evident that

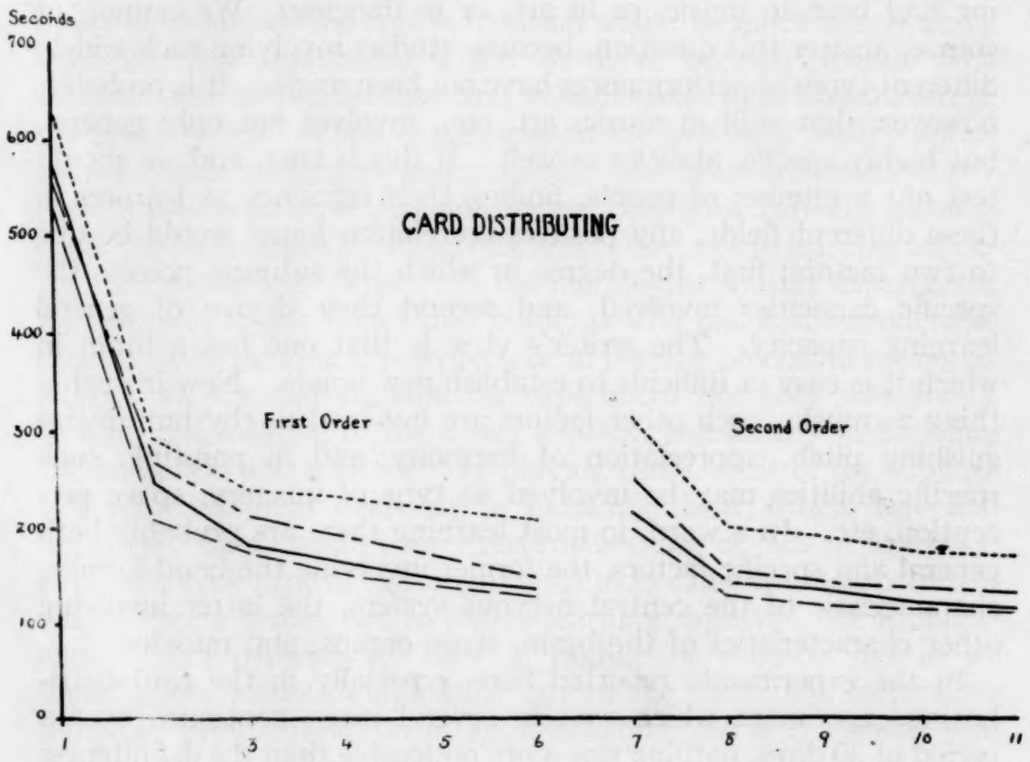
if these extraneous factors could be eliminated, the true correlation would be much higher. In the writer's opinion it would be unity. But of course this has not been demonstrated. It remains to eliminate all the other factors and find the relation of learning capacity itself in one kind of work with learning capacity in another.

However, there is another possibility. It may be that our assumption of a definite learning capacity in an individual for all types of material is erroneous, and that the positive correlation which we have found is due to the supposed "central" or "general" factor of the English school. The central factor may be *one* element in producing the positive correlation, but it can scarcely be the *sole* element. The inter-correlation in the following tests,—logical memory, rote memory, substitution, opposites, free association, word-building, and completion,—averages .231 as found by the writer in a careful study of sixty cases. If there is such a thing as central factor it certainly operates in all these seven tests. Still the inter-correlation is less than half that which we have found in the case of the four different types of learning.

How would this experiment have turned out if some of the learning had been in music, or in art, or in dancing? We cannot, of course, answer this question, because studies involving such widely different types of performances have not been made. It is probable, however, that skill in music, art, etc., involves not only general, but highly specific abilities as well. If this is true, and we should test out a number of people, finding their efficiency as learners in these different fields, any positive correlation found would be due to two factors; first, the degree in which the subjects possess the specific capacities involved, and second their degree of general learning capacity. The writer's view is that one has a brain in which it is easy or difficult to establish new bonds. Now in such a thing as music such other factors are involved as rhythm, distinguishing pitch, appreciation of harmony; and in painting, such specific abilities may be involved as type of imagery, space perception, etc. In a word, in most learning there are probably both general and specific factors, the former involving the bond-forming characteristic of the central nervous system, the latter involving other characteristics of the brain, sense organs, and muscles.

In the experiments reported here, especially in the card-distributing experiment which was, in several cases, continued over a period of 30 days, nothing was more noticeable than the definiteness

and regularity of the learning of the different subjects. They retained their relative positions or ranks from day to day with great constancy as, for example, in the card-distributing experiment as shown in the accompanying graph. At the beginning of an experiment the experimenter could predict with great accuracy the relative ranks of the subjects on the basis of their work on the previous day. The learning efficiency of the various subjects became definitely known to the experimenter. The characteristics underlying learning efficiency seem to persist and to be as definite as anything in nature. Such variations as occur in a given subject's learning from day to day have definite causes, and can usually be predicted when the causes are known to exist. Of course, constancy in learning always depends upon constancy of the conditions and factors involved. These factors are numerous. The condition of nearly every organ in the body can have its effect upon the temporary efficiency of the learner. But there is probably a constant factor which may be called general learning capacity, dependent upon the characteristics of the central nervous system.



MENTAL TESTS FOR COLLEGE ENTRANCE*

L. L. THURSTONE

Carnegie Institute of Technology, Pittsburgh

PURPOSE OF THE MENTAL TESTS

One of the main duties of the division of applied psychology at Carnegie Institute of Technology is to advise the deans and the faculty concerning the admission of students and their subsequent promotion and transfer from one course of study to another. Our main concern so far has been to establish the diagnostic value of mental tests as a criterion for admission. In order to demonstrate the practical value of mental tests we have been giving a series of tests to the freshmen early in their first year of residence. The test papers were scored and filed until scholarship records and faculty estimates were available. Then we asked ourselves this question: If we had been given complete control over admissions could we have predicted by means of our tests which students were to fail and which were to succeed in their college work? This is a practical problem. Our results may be ever so interesting from some theoretical standpoint but even if we compile tons of norms of performance, the work is a failure if we are unable to establish a relation between the test scores and the criterion which is in this case the student's performance in college work. The present report is an answer to this question with reference to the entering students at the women's college at Carnegie Institute of Technology.

DESCRIPTION OF GROUP

The students to whom we gave these tests were all freshmen at the Margaret Morrison Carnegie School of Carnegie Institute of Technology in Pittsburgh. All of the 114 freshmen were high school graduates. The four year course in the women's college is devoted mainly to applied science, including household economics, secretarial studies, costume economics, home arts and crafts, social work and general science. The freshmen look forward to specializing in one of the above mentioned departments. For this reason it is apparent that the school does not attract students of academic preferences, but the grade of work in the special subjects mentioned above is comparable with that of other well known women's colleges. The first year course, on which the instructors based their estimates of the students, contains the following subjects: physics, sewing, history, English, drawing and color, hygiene, chemistry, foods, accounting, and social ethics.

*Since this article was written our mental test program has been arranged in the cycle-omnibus form, using most of the test material here described. The results of the cycle-omnibus form will be published later.

Description of the Tests

I shall limit myself in this report to six of the mental tests which constitute our present standard series for the women's college at the Institute. These tests are as follows:

1. Completion Test

This is a selection from the Trabue Language Test material used by Professor Walter Dill Scott in the Bureau of Salesmanship Research at the Carnegie Institute of Technology.

2. Robinson Range of Interest Test

The instructions for the test are as follows:

Below is a list of statements. Some of these statements are true: some are false. In the ( ) to the left of each statement make a + sign if the statement is true, but a − sign if the statement is false. Even when in doubt make the mark which you think is most probably correct. Be sure to mark each statement. The first two statements are already marked correctly."

(+) 1. George Washington was the first president of the United States.
(−) 2. Christmas comes on July 4.

The following ten sample sentences illustrate the nature of the test.

( ) 3. A satisfactory rating in "Bradstreet's" is an important business asset.
( ) 8. A broken crank shaft generally means a ruined carburetor.
( ) 19. Arson is sometimes used as a medicine.
( ) 23. Sir Roger Casement was accused of treason and put to death.
( ) 49. Forbes Robertson's performance of "Hamlet" has thrilled many audiences.
( ) 51. The "squeeze play" calls for a long hit by the batter.
( ) 52. Stocks and bonds are sometimes offered as collateral.
( ) 69. The Iroquois fire would have been less serious if there had been an adequate supply of lifeboats.
( ) 101. John Marshall was a famous American jurist.

3. Opposites Test

This test has the following instructions:

"On the following page are forty pairs of words. If the two words of a pair mean the same, or nearly the same, draw a line under the word SAME. If they mean the opposite, or nearly the opposite, draw a line under the word OPPOSITE. If you cannot be sure, guess. The two samples are already marked as they should be."

fall–rise	same	*opposite*
class-group	*same*	opposite

"When the signal is given (wait for the signal) turn over this page and work as rapidly as you can without making mistakes."

The test proper has forty reactions like the samples in the instructions.

4. Gordon Directions Test

This test has four parts printed on the four pages of a pamphlet. Each part of the test has forty stimulus words. The instructions are as follows:

"Part 1. Work as quickly as you can without making mistakes. At the right of each of the following words make a figure. If a word contains the letter **e** mark it 1. If it does not contain **e** mark it 2. Look at the samples. Samples: Strike 1, train 2."

"Part 2: Work as quickly as you can without making mistakes. At the right of each of the following words make a figure. If a word contains the letter e and a mark it 1. If it contains the letter **e** but not **a** mark it 2. If it contains the letter a but not e mark it 3. Samples: change 1, every 2, calamity 3."

"Part 3: Work as quickly as you can without making mistakes. At the right of each of the following words make a figure. If a word contains the letters **e**, **a**, and **r**, mark it 1. If it contains the letter **e** but not **a** and **r**, mark it 2. If it contains the letter **a** but not **e** and **r** mark it 3. If it contains the letter **r** but not **e** mark it 4. Samples: dream 1, feeling 2, stand 3, stirrup 4."

"Part 4: Work as quickly as you can without making mistakes. At the right of each of the following words make a figure. If a word contains the letters **i**, **n** and **g**, mark it 1. If it contains **i** and **n** but not **g** mark it 2. If it contains **i** and **g** but not **n** mark it 3. If it contains **n** and **g** but not **i** mark it 4. If it contains i but not n and g mark it 5. Samples: ignite 1, thirteen 2, ligature 3, gnarled 4, ritual 5."

5. Analogies Test

The instructions for this test are as follows:

"On the following pages are forty lines of words. In each line the first two words are related in a certain way. Draw a line under the word in heavy type which has the same relation to the third that the second has to the first.

Notice the following samples:"

1. Foot-shoe; hand (head, *glove*, finger, clasp)
2. Mayor-city: captain (*ship*, private, general, town)
3. Rudder-ship: tail (*dog*, cat, sail, *bird*)
4. Penny-copper: dime (quarter, *silver*, nickel, spend)

The test proper contains forty problems.

6. Marble Statue Test

This memory test was given and scored according to instructions in Whipple's *Manual of Mental Tests*. Vol. II., p. 209.

Procedure in Giving the Tests

All the tests are printed in uniform size, 8½"x11." They are bound into a pamphlet, stitched and cut. One pamphlet is given to each subject. The tests are given to groups of about fifty students simultaneously. The instructions for the test are printed on one page, followed on the next page by the test itself. This enables the examiner to control the amount of time allotted to reading the instructions and the amount of time for the test proper. The Gordon Directions Test has the instructions printed on the

same page with the test itself since the test score is intended to reveal the subject's alertness in comprehending the instructions. The time limits for the tests are as follows:

1. Completion test. 10 minutes.
   Special scoring method for each response. From two to four points are allowed for each perfect sentence.
2. Robinson Range of Interest Test. 20 minutes.
   This is ample time for all subjects to finish the test. The test is really given by the amount limit method since every candidate is expected to finish the whole test. Scoring formula: $S = 100 - 2W - O$ in which
   S = Total Score
   W = Number of wrong responses
   O = Number of omitted responses
3. Gordon Directions Test. Part 1, one minute; Part 2, one and one-half minutes; Part 3, two minutes; Part 4, two and one-half minutes. The score is the total number of right responses on the last three pages.
4. Analogies Test. Two minutes.
   Scoring formula: $S = 100 + R - 3W$ with the same notation as in the Range of Interest Test.
5. Marble Statue Test. This test is read by the examiner and the subjects are asked to reproduce the paragraph after an interval of thirty minutes filled with other tests. The score is the number of ideas reproduced. Just what constitutes an idea is left to the interpretation of the scorer with the help of the paragraph dissected more or less arbitrarily into ideas.
6. Opposites Test. Two minutes.
   Scoring formula: $S = R - W$ with notation as before.

The Criterion

The value of a mental test is determined solely by its agreement with a criterion. Much has been said about scholarship as a criterion. If our task is to ascertain whether or not a candidate is intellectually able to do college work, and if the students' retention, promotion, dismissal, or graduation is determined by his scholarship grades, then there can be no other criterion than scholarship. That seems so obvious that there is really no room for debate about the question. To use any other criterion constitutes an indictment against the scholarship marks as a criterion for retention and promotion of college students. If our tests agree well with estimated intelligence but fail to agree with faculty action concerning the dismissal of poor students and the promotion of good students, then

our tests are of little use to the administrative officers of the college, unless they are willing to be guided by our statement of the student's intelligence rather than their own scholarship ratings.

In view of what has been said *pro* and *con* regarding scholarship as a criterion for the efficiency of mental tests we have used both criteria. Our present series of six tests have been evaluated against the following criteria:

(1) Can we demonstrate that we can reduce the number of students who are dropped for poor scholarship or placed on probation for poor scholarship by the use of our mental tests for admission?

(2) How do our mental test ratings of all the students compare with the faculty opinion about the general ability of the students?

The first criterion refers only to those who were pronounced as failures and dropped from college for inability to do college work or placed on probation as doubtful students with two thirds of the regular program. The second criterion has reference to the whole class including the good students.

The second criterion by which our tests have been evaluated is the median instructor's estimate of the student's general ability. We sent the following letter to all the instructors who have freshmen in their classes:

> *To the Members of the Faculty:*
>
> In order to evaluate the psychological tests which were given to the freshmen recently it is necessary to ask for your co-operation in judging the students. Those tests which agree fairly well with the pooled judgment of the faculty will be retained and the tests which fail in this regard will either be improved or cancelled.
>
> Kindly place a check mark in one of the ten spaces after each name to indicate the student's GENERAL ABILITY compared with the FRESHMAN CLASS. General ability is to be thought of as broader than mere scholarship. Leave unmarked the names of students you do not know or whom you have not had in your class for at least one semester. Try to place approximately the same number of students in each of the ten classes.
>
> Please return this report to the Dean's office by ————.

Attached to the letter was a list of the names of the Freshmen with ten numbered spaces after each name. When the returns were complete the median instructor's estimate was determined for each student and this was used as a criterion for the tests. These instructors' judgments were given in March, 1918. The tests were given in October, 1917.

Method of Evaluating the Mental Tests

There are two fundamentally distinct problems connected with the evaluation of a mental test. These are (1) Is the test diagnostically significant when applied to each individual student? and (2) What is the degree of relationship between the two variables without reference to any individual subject? These problems are restatements of the two problems in the preceding section.

The Pearson coefficient of correlation and the correlation ratio express the degree of relationship between the variables as such and the regression equation can be used for predicting one of the variables when the other is known. By means of the standard error of estimates we may quantify the reliability of the prediction for any individual but this evidence is not as satisfying as when the results are stated in terms that bring out the individual members of the group. The combination of tests can be handled by multiple correlation but the procedure is in practice hardly worth the trouble if our problem centers on the diagnostic value of the test for the individual rather than for the group. However, in order to extract all possible information from our tests they have been analyzed both by correlation methods referring to the group as a whole, and by inspection of the scatter diagrams referring to the individual student. I shall refer to the latter procedure as *the method of critical scores.*

The Method of Criterial Scores

The median instructors' estimate is ascertained for each student. This is based on the opinion of from three to six instructors. A scatter diagram is plotted for each test. Figure 1 is the scatter diagram for the Trabue Completion Test. The Pearson coefficient of correlation between the pooled estimate of the instructors and the score in this test is $+.49$.

The next step is to determine an upper critical score and a lower critical score. These are indicated in figure 1. They are so drawn that every student below the lower critical score is below the mean of the instructors' estimate. Every student above the upper critical score is above the average in the opinion of the faculty. Some tests have fair correlation coefficients but do not yield any critical scores on account of exceptions at the extremes. Some mental tests yield only one critical score. Thus when a test gives an upper critical score but no lower critical score we conclude that the test is useful for selecting unusually good students but not for spotting poor students.

The reader will notice that the lower critical score for the Trabue Completion Test is at 48 and that all the nine students who scored below 48 were rated below the average by the faculty. Facts of this type are practically useful, and when combined with similar facts from other tests, they are diagnostically significant in qualifying and disqualifying students for college work.

Notice also that the upper critical score for the Trabue Completion Test (Figure 1) is at 84 and that all the students who scored above 84 in this test are rated above the average by the faculty. The horizontal double line on the chart indicates the mean of the faculty estimates.

When the critical scores have been ascertained for all tests, they are combined in order to enhance the reliability of the prediction. This is done as follows: On the student's mental test record card each score that is above its upper critical score is marked in red, and each test score that is below its lower critical score is marked in blue. The median percentile rank in all tests is also recorded. Then we obtain a list of all the students who have been dropped during the year for poor scholarship and also those who have been placed on probation or two-thirds schedule on account of poor scholarship. Our motive is to obtain a method of predicting these failures by our test records. In order to do this we prepare lists of students for conditions such as the following:

(1) A list of all students whose median percentile rank in six tests is 20% or below, 15% or below, 25% or below, etc.

(2) A list of all students whose median percentile rank in six tests is 20% or below, and who were below the lower critical score in two or more of the tests, three or more of the tests, one or more of the tests, etc.

(3) A list of all students whose median percentile rank in the six tests is 20% or below, and who were below the lower critical score in more tests than they were above the upper critical score.

The lists so provided are compared with the list of failures due to poor scholarship. We select as our standard method that which predicts the largest proportion of failures *without excluding any students who have made good.* This method intensifies the responsibility of the psychologist with regard to the individual. The correlation methods do not place this emphasis on the individual but state the abstract relation between the several variables and this does not bring home to the psychologist his chances of error in advising about an individual student.

After an extensive analysis of our test records and the list of failures we found that the following criterion was quite satisfactory:

Those students are reported as doubtful who attain a median percentile rank of 22% or below, and who are below the lower critical score in two or more of the six tests.

The Mental Test Rating

In order to quantify this test analysis into a scale we have adopted the procedure of designating as the mental test rating the median percentile rank in all six tests plus five points for each test in which the student is above the upper critical score and deducting five points for each test in which he is below the lower critical score. We report as doubtful all students whose mental test rating is +10 or below.

By inspecting Figure 2 the reader will notice that we could have eliminated seven out of the eleven total failures at the beginning of the year. Instead of eleven total failures there would have been only four if the mental test rating had been a criterion of college entrance. Furthermore, we should have eliminated at the beginning of the year eight of the seventeen students who were placed on probation for poor scholarship. Instead of seventeen such cases there would have been only nine cases of probation for poor scholarship. Another very important fact is that not one of the students who were below the lower critical mental test rating was acceptable as a student. All of them without exception were either dropped for poor scholarship or placed on probation with reduced program on account of poor scholarship. They should have been saved from the discouragement which comes with failure and should have been advised to take up some other work. None of the acceptable students scored below the lower critical mental test rating.

Figure 2 also shows that the eleven students who were rated as unusually good students by the faculty were all above the average according to the mental test rating.

In figure 3 we have the scatter diagram for instructors' estimate against the combined mental test rating.

The Pearson correlation coefficient for these two variables is +.60. It is interesting to know that the application of the method of multiple correlation to these two variables for a closely parallel set of six tests gives a total coefficient, R, of +.67 which is not alarmingly superior to the method of critical scores and certainly

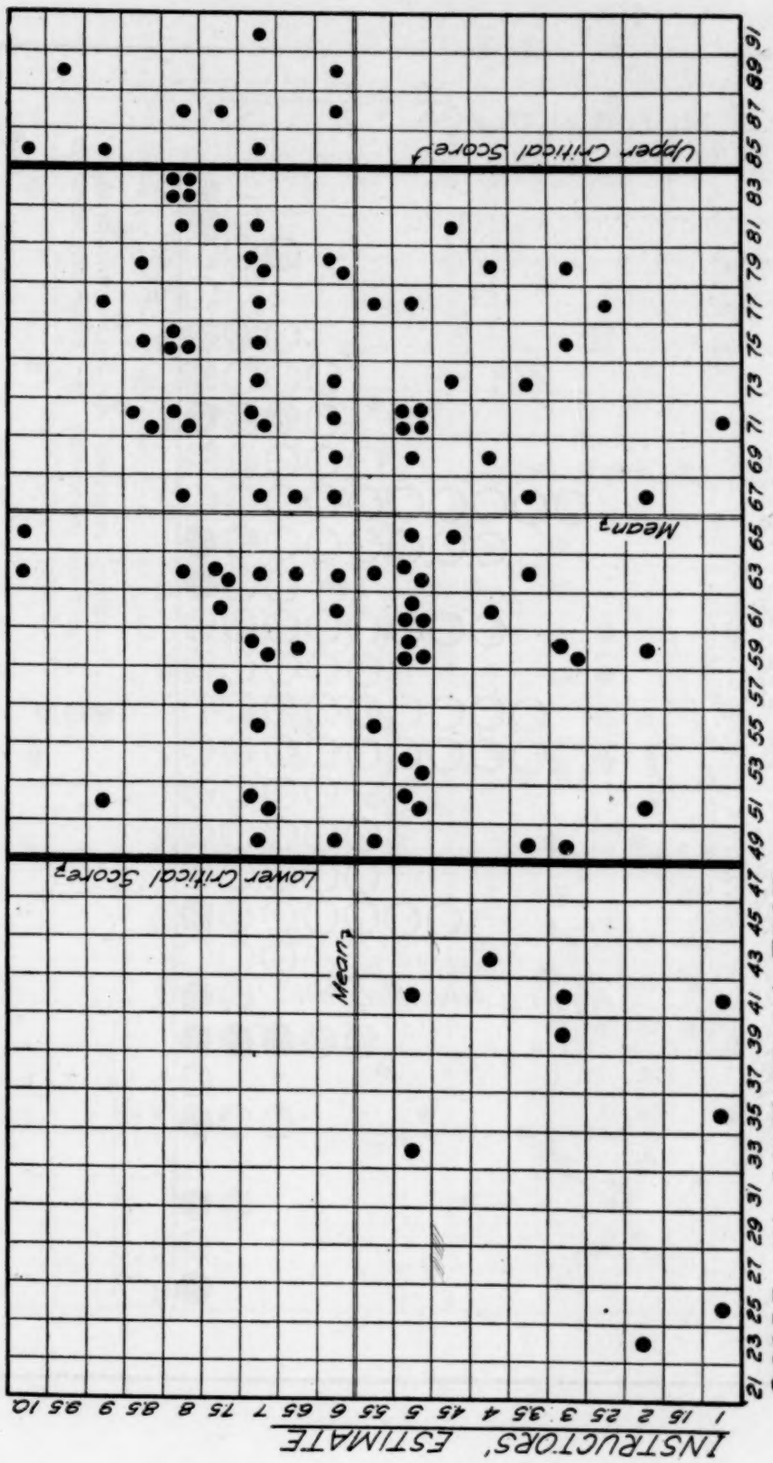


FIGURE 1.

Scatter diagram showing relation between median instructors' estimate and score in completion test. After the means have been drawn the two critical scores are so selected that all students above the upper critical score are above the average in the opinion of the faculty and all students below the lower critical score are below the average in the opinion of the faculty.

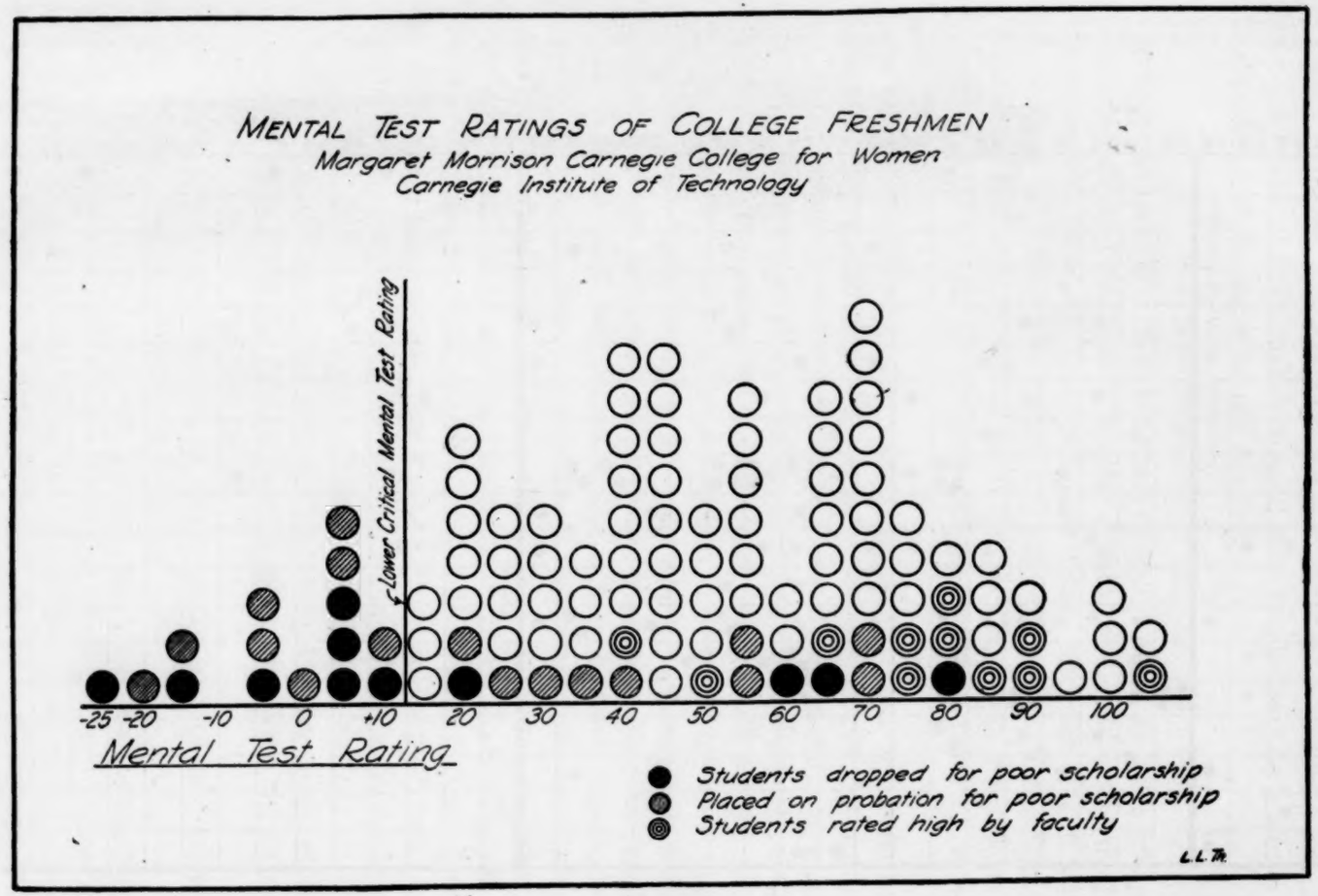

MENTAL TEST RATINGS OF COLLEGE FRESHMEN
Margaret Morrison Carnegie College for Women
Carnegie Institute of Technology
Lower Critical Mental Test Rating
-25 -20 -10 0 +10 20 30 40 50 60 70 80 90 100
Mental Test Rating
Students dropped for poor scholarship
Placed on probation for poor scholarship
Students rated high by faculty
L.L.Th.

FIGURE 2.

Frequency distribution of mental test rating which is the median percentile rank in six mental tests adjusted by adding five points for each test in which the candidate is above the upper critical score, and deducting five points for each test in which the candidate is below the lower critical score. Note that seven of the eleven freshman total failures were below the lower critical mental test rating.

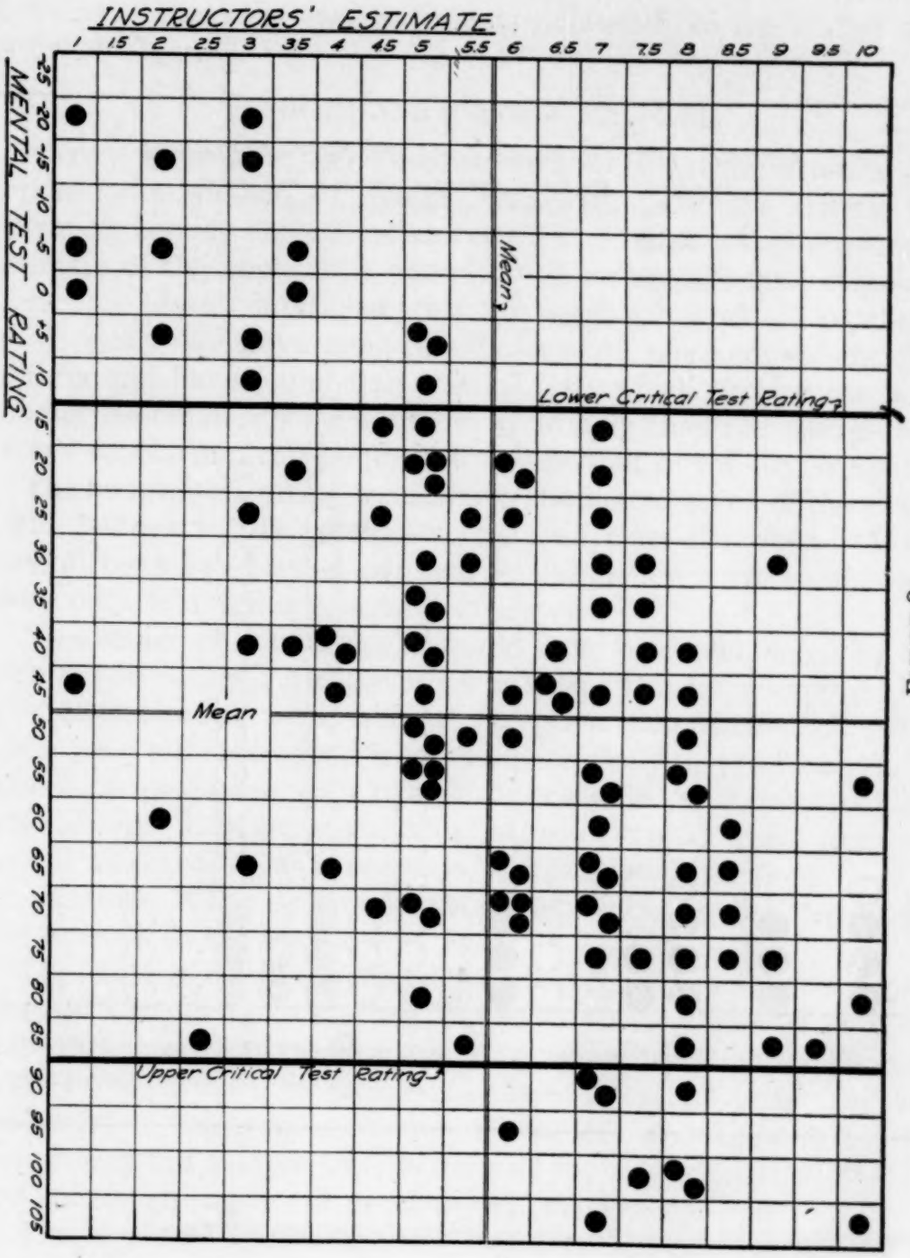

INSTRUCTORS' ESTIMATE
1 1.5 2 2.5 3 3.5 4 4.5 5 5.5 6 6.5 7 7.5 8 8.5 9 9.5 10
MENTAL TEST RATING
-25 -20 -15 -10 -5 0 +5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 105
Mean
Lower Critical Test Rating
Mean
Upper Critical Test Rating

FIGURE 3.

Scatter diagram showing relation between instructors' estimate and mental test rating. Note that all the freshmen rated below the lower critical mental test rating (+10.) are below the average in the opinion of the faculty and that all who scored above the upper critical mental test rating (85) are rated above the average in the opinion of the faculty.

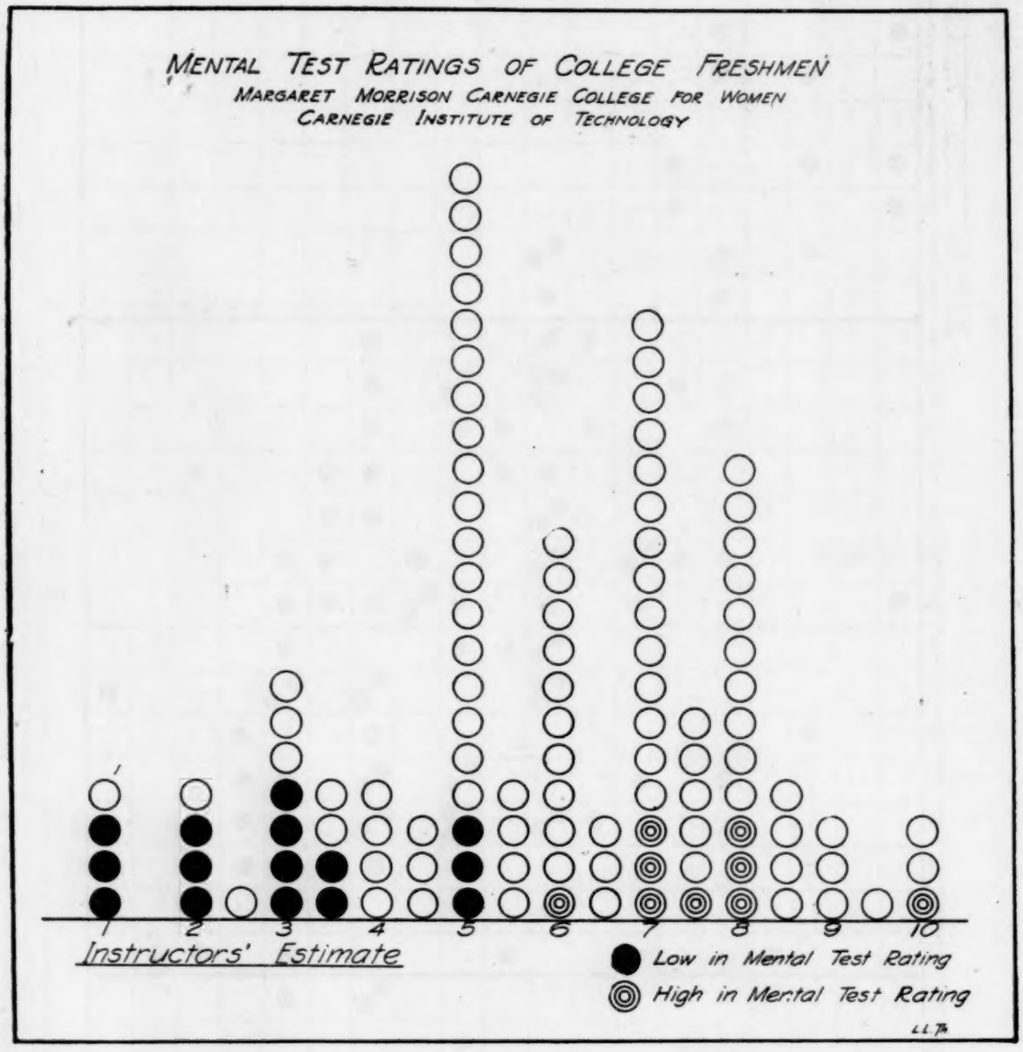


FIGURE 4.
Frequency distribution of instructors' estimate showing the same facts as in figure 3.

not nearly as controllable. Figure 3 shows that all students below the lower critical mental test rating were rated below the average by the faculty and that all students above the upper critical mental test rating were rated above the average by the faculty. In figure 4 we have these last two statements isolated from the scatter diagram of Figure 3.

Summary and Recommendations

Mental tests have been demonstrated to constitute a useful criterion for admission to college. We shall have five criteria for entrance, namely (1) high school credits, (2) high school principal's estimate of the applicant's capacity, (3) college entrance examinations, (4) Dean's interview with the applicant, and (5) psychological tests. We have analyzed the relative efficiency of the five criteria and have found that mental tests which occupy one hour predict freshman scholarship more accurately than the second semester's scholarship can be predicted from that of the first semester.

The mental test rating would have eliminated seven of the eleven total failures in the particular freshman class described. If the mental tests had been a criterion for admission there would have been only four failures instead of eleven.

No average or good student would have been eliminated by the mental test rating. All students who scored below the lower critical mental test rating were, without exception, poor students. All of them were either dropped from college entirely in the first year or placed on probation with two-thirds schedule on account of poor scholarship. Hence no individual injustice would result from the use of mental tests as a criterion of college entrance.

All of the freshmen who were rated high by the faculty were above the average in the mental test rating.

The standard series of six tests has been given to the freshman class of 1918–19 but in addition four new tests have been given which are in the process of being accepted or rejected. If any of the new tests prove to be more successful than last year's tests they will be added to the standard series and the old tests discarded. By continuing this process the efficiency of the test series as a criterion for college entrance will probably be raised every year.

The method of critical scores has been applied successfully in standardizing the Army Trade Tests. The method is more controllable than the method of multiple correlation for establishing diagnostically significant criteria. One of its main advantages is

in the ease of application. It emphasizes prognosis with respect to each individual student rather than the abstract statement of relationship between the variables concerned.

The method of critical scores consists simply in plotting a scatter diagram for each test, showing relation between scholastic attainment and mental test score for each individual. An upper critical score is selected such that all who score above it are above the average in scholastic attainment. A lower critical score is selected such that all who score below it are below the average in scholastic attainment. Some mental tests do not yield any critical scores.

When tests are combined by the method of critical scores the median percentile rank is corrected by adding five points for each test in which the subject is above the upper critical score and deducting five points for each test in which the subject is below the lower critical score.

This emphasizes the extremes. The same effect of emphasizing the extremes is obtained by using the standard deviation but it is not so readily interpreted. I do not wish to discourage the use of multiple correlation as a means of analysis but I do insist that we should not be enslaved by its laborious refinements in making mental test results practically applicable.

THE NECESSITY OF TEACHING DERIVED FORMS IN SPELLING

E. HORN AND E. J. ASHBAUGH

University of Iowa

Introduction. Progress in the scientific study of spelling has been made along three major lines: first, in selecting and grading the words which are to make up the course of study; second, in the experimental determination of the most economical methods of learning; and, third, in the development of satisfactory scales for the measuring of spelling ability. In all of these lines research has proceeded well beyond the pioneer stage so that we may now expect a series of studies which will greatly refine our present knowledge and practice. The investigation described in this article was undertaken in the attempt to accomplish one such refinement.

From the beginning of vocabulary studies in spelling, the writers believed that each change in the form of a word should be scored as a separate word, and that attempts to group according to root forms were objectionable. Particularly objectionable is the practice of deciding, without data, whether or not a given form constitutes special spelling difficulty. The following statement from Ayres, *Measuring Ability in Spelling*, illustrates the use of judgment in determining whether or not a given form is difficult to spell.

"In making up this list, there has been no attempt to reduce all the words to a dictionary basis. Instead the attempt has been to include all the forms of the words which present different spelling difficulties. Thus the various forms of the verb "be" are included as separate words because they present separate spelling difficulties. In the same way, "man" and "men," "woman" and "women," are included for the same reason. On the other hand, plurals and verb forms presenting no characteristic spelling difficulties beyond those inherent in the singular or infinitive have not been included. This procedure has necessitated making many arbitrary decisions, but in each case the controlling purpose has been to make each decision on the basis of spelling difficulty."

The data which follow show conclusively that in the great majority of cases a change in form is accompanied by a difference in spelling difficulty, and that with few exceptions derived forms are more difficult than the roots upon which they are built. According-

ly it becomes important to know the spelling difficulty of each derivation and also the frequency with which it occurs in correspondence.

All data in Tables I, III, IV, V, and VI are based upon the spelling of a random selection of two hundred children from the second grades in the schools of Iowa. The words chosen for Table I were those root forms upon which the children spelled with an accuracy of 60 percent. or better, and for which the accuracy on one or more derived forms was also available. The words in Tables III, IV, V, and VI were chosen by taking in regular order from the alphabetized list words appearing in both root form and with the suffix under consideration.

Table I shows the accuracy of spelling of second grade children upon the root and derived forms of forty words. The table is read as follows: The accuracy of spelling of second grade children on the word *be* was 89 percent. and on the derived form, *being*, was 65 percent.,—a difference of 24 percent. In all cases in this and the following tables, when the accuracy on the derived form was less than on the root form, the difference is marked minus; when the accuracy on the derived form was greater than on the root form, it is marked plus.

The seventh word has two derived forms and is read thus: The accuracy of second grade children on the word *look* was 82 percent., on *looking* was 70 percent., and on *looked* 40 percent. The difference between the accuracy on the root form and the first derived form was 12 percent. and between the root and second derived form was 42 percent. The accuracies and differences on all other words in the table are read in the same manner.

The last column at the right is headed "Group Average." Under this is shown for the first and second groups of twenty words separately the average difference between the accuracy on the root form and on the first, second and third derived forms. It will be noted that the average difference in accuracy between the root form and the derived form is greater in the group with the higher accuracy on the root form.

At the foot of the table is given the averages and medians for accuracies and differences on the entire list of forty words. The decrease in accuracy of spelling on the derived forms is really greater than appears in the average differences since they must be thought of in terms of the accuracy on the root forms upon which they are

TABLE I

Difference between root and derived form in accuracy of spelling.

root	root acc.	Accuracy of derived forms 1st		2nd		3rd		Difference from root Acc. 1st	2nd	3rd	Group Ave.
be	89	ing	65					−24			
go	89	ing	68					−21			
see	88	ing	71					−17			
lone	84	ly	14					−70			
come	83	ing	46					−37			
do	83	ing	72					−11			
look	82	ing	70	ed	40			−12	−42		1st
tell	82	ing	64					−18			−32.65
have	81	ing	39					−42			
will	81	ing	44					−37			
big	80	ger	23	gest	5			−57	−75		
land	80	ing	58					−22			2nd
old	80	er	42	est	36			−38	−44		−51.8
say	80	ing	52					−28			
hand	79	ed	43					−36			
run	77	ning	9					−68			
dear	76	ly	29	est	29			−47	−47		
play	76	ing	68	ed	25			− 8	−51		
cold	73	er	42					−31			
think	73	ing	44					−29			
bad	70	ly	37					−33			
love	69	d	40	ly	40	ing	15	−29	−29	−54	
box	68	es	34					−34			
like	68	ly	32	d	17			−36	−51		
call	67	ing	45	ed	33			−22	−34		
put	66	ting	11					−55			
fast	65	er	49					−16			
fill	65	ing	34	ed	21			−31	−44		
find	65	ing	54					−11			
glad	64	ly	32					−32			1st
name	63	d	29	ly	22			−34	−41		−26.55
bill	62	ed	14					−48			
kiss	61	es	24					−37			
part	61	ly	24					−37			2nd
send	61	ing	35					−26			−34.56
want	61	ing	49	ed	26			−12	−35		
hold	60	ing	55	er	33			− 5	−27		
live	60	d	46	s	43	ing	22	−14	−17	−38	3rd
stand	60	ing	60					0			−45.3
wish	60	ing	41	es	27	ed	16	−19	−33	−44	
Ave.	72.3	42.7		28.4		17.7		−29.6	−40.7	−45.3	
Median	71.5	43.0		28.0		16.0		−30.0	−41.5	−44.	

SPELLING

TABLE II

Difference between root and derived form in frequency of occurence. Grade II

	Frequency of derived forms								
root		1st		2nd		3rd		4th	
be	4042	ing	305						
go	656	ing	640						
see	824	ing	14						
lone	16	ly	12						
come	882	ing	312						
do	1309	ing	235						
look	320	ing	109	ed	50				
tell	359	ing	20						
have	4781	ing	331						
will	4211	ing	68						
big	278	ger	19	gest	15				
land	249	ing	13						
old	215	er	71	est	8				
say	666	ing	43						
hand	308	ed	31						
run	111	ning	32						
dear	2194	ly	7	est	21				
play	84	ing	15	ed	23				
cold	220	er	5						
think	650	ing	54						
love	282	d	8	ly	51	ing	19		
box	120	es	24						
like	744	ly	15	ed	33				
call	334	ing	50	ed	123				
put	243	ting	40						
fast	28	er	8						
fill	102	ing	26	ed	40				
find	510	ing	17						
glad	760	ly	42						
name	423	ly	9	d	15				
bill	170	ed	6						
kiss	8	es	13						
part	217	ly	12						
send	657	ing	194						
want	583	ing	19	ed	118				
hold	157	ing	28	er	25				
live	113	d	15	s	21	ing	47		
stand	68	ing	20						
wish	539	ing	43	es	205	ed	66		

based. For example, the average difference of the first derived forms is 29.6 percent. Since the average accuracy on the root forms was 72.3 percent. this 29.6 percent. represents a decrease of 40.9 percent. The average difference of the second derived forms is 40.7. But since only fourteen words are represented by second derived forms, this must be considered in terms of the accuracy of these fourteen words or 69.1 percent. and hence represents a decrease of 58.9 percent. In like manner the average of the three third derivatives represents a decrease of 71.9 percent.

Table II gives for each root and derived form in Table I the frequency of occurance in two studies in which the words have been scored separately. One of these studies,[1] represents an analysis of 1175 letters of bankers to bankers; the other,[2] an analysis of 3763 letters written by Iowa correspondents. These two studies represent the scoring of over 400,000 running words.

The tables should be read: *be*, has a frequency in these combined lists of 4042; *being*, a frequency of 305, etc. The only case in which the derivation occurs more frequently than the root form is in the word kiss, which is exceeded in frequency by kisses.

Table III represents the difference between the accuracy of second grade children in spelling root forms and the same words with the suffix *ing*. The first twenty words attach the suffix without any change in the form of the root word and the second twenty drop the final *e* before adding *ing*. The table is read as follows: The accuracy of second grade children on the word *ask* was 54 percent. and on the derived form, *asking* was 42 percent. Hence the difference between the two forms is 12 percent. The percentages on the other words are read in the same manner. Differences showing the derived form less than the root are proceded by a minus sign, and those showing the derived form greater, by a plus sign.

The table has been divided horizontally to indicate groups of words upon which the accuracy on the root forms fell within the same quartile percentages, that is 100 percent.—76 percent.; 75 percent.;—51 percent.; 50 percent.—26 percent, and 25 percent.—

---

[1] ERNEST HORN. *A Study of the Vocabulary of Letters Written by Bankers to Bankers about Banking.* Study unpublished.

[2] W. N. ANDERSON. *The Determination of a Spelling Vocabulary Based upon Written Correspondence.* Ph.D. Thesis, University of Iowa, 1917.

TABLE III

Difference between root and derived form. Grade II.

Word	Freq. Root	Freq. Der.	Acc. Root	Acc. Deriv.	Dif. in Acc.	Aver. dif. in group
ing–						
ask	263	67	54	42	−12	
bank	569	299	53	30	−23	
belong	17	7	53	32	−21	
bring	128	19	59	49	−10	
call	334	50	67	45	−22	−18
act	56	7	31	17	−14	
add	77	6	28	32	+ 4	
amount	272	33	29	5	−24	
buy	3	51	40	20	−20	
check	224	31	33	12	−21	
clean	55	19	35	27	− 8	−14
answer	250	123	10	10	0	
attend	74	19	20	28	+ 8	
await	13	69	16	15	− 1	
board	129	6	22	9	−13	
break	21	9	7	12	+ 5	
build	41	110	12	15	+ 3	
carry	86	33	23	27	+ 4	
coast	17	5	10	12	+ 2	
collect	36	9	6	0	− 6	+0.2
come	882	312	83	46	−37	
have	4781	331	81	39	−42	−40
bake	15	6	48	31	−17	
change	298	16	30	14	−16	
inclose	15	12	38	25	−13	
leave	143	31	39	10	−29	−19
advertise	18	50	8	9	+ 1	
advise	242	34	16	10	− 6	
anticipate	13	6	0	2	+ 2	
appreciate	308	10	1	1	0	
arrive	37	7	4	1	− 3	
assure	94	70	2	4	+ 2	
believe	295	11	8	6	− 2	
choose	22	8	19	14	− 5	
dance	21	11	20	16	− 4	
debate		5	20	17	− 3	
desire	192	7	12	15	+ 3	
figure	59	11	8	3	− 5	
increase	54	28	4	7	+ 3	
invite	1	29	18	23	+ 5	− 0.9

TABLE IV

Difference between root and derived form. Grade II.

-ed	Word	Freq. root	Freq. Der.	Acc. Root	Acc. Deriv.	Dif. in Acc.	Aver-dif. in group
	ask	263	122	54	31	−23	
	bill	170	6	62	14	−48	
	call	334	123	67	33	−34	−35
	act	56	10	31	7	−24	
	add	77	44	28	42	+14	
	check	224	6	33	7	−26	
	clean	55	29	35	14	−21	
	cover	121	31	35	14	−21	−16
	accomplish	14	6	3	1	− 2	
	address	224	100	19	8	−11	
	adjust	12	14	22	7	−15	
	appear	36	18	6	1	− 5	
	attach	10	22	9	9	0	
	attend	74	13	20	3	−17	
	allow	53	26	10	6	− 4	
	answer	250	40	10	12	+ 2	
	connect	19	21	1	2	+ 1	
	consider	164	49	7	5	− 2	
	contain	30	12	9	12	+ 3	
	correct	54	7	10	1	− 9	− 5
-d	agree	421	8	32	20	−12	
	care	283	21	40	27	−13	
	change	298	45	30	14	−16	
	charge	230	1	30	13	−13	−13.5
	advance	88	45	9	1	− 8	
	advertise	18	15	8	5	− 3	
	advise	242	26	16	4	−12	
	appreciate	308	52	1	3	+ 2	
	arrange	67	21	3	3	0	
	arrive	37	48	4	4	0	
	associate	5	7	0	1	+ 1	
	assure	94	25	2	2	0	
	cause	57	22	11	11	0	
	continue	70	32	9	8	− 1	
	convince	15	9	9	5	− 4	
	decide	91	82	6	6	0	
	deserve	8	6	5	2	− 3	
	desire	192	38	12	7	− 5	
	dispose	9	14	18	22	+ 4	
	figure	59	14	8	6	− 2	− 2

TABLE V

Difference between root and derived form of words adding er. Grade II

Word	Freq. Root	Freq. Der.	Acc. Root	Acc. Deriv.	Dif. in Acc.	Aver. dif. in group
old	215	71	80	42	−38	
bank	569	25	53	45	− 8	
cold	220	5	73	42	−31	
cool	28	7	52	35	−17	
farm	186	92	54	36	−18	
feed	27	9	54	39	−15	
hold	157	25	60	33	−27	
kind	507	6	51	35	−16	
long	374	60	54	40	−14	
low	71	28	52	16	−36	
near	182	9	52	31	−21	−20.3
clean	55	10	35	27	− 8	
high	280	37	26	24	− 2	
own	243	63	36	5	−31	
paint	57	8	27	15	−12	
print	23	7	50	29	−21	
rent	63	9	46	33	−13	−14.5
build	41	7	12	8	− 4	
cheap	31	12	10	9	− 1	
commission	76	7	2	2	0	− 1.7

0 percent. The average differences within these groups are shown to the right and the same fact is noted as in Table I, namely, that the average differences are greater in groups having the higher accuracy upon the root forms.

Table IV presents the same kind of data upon forty words with the suffix *ed* as is given in Table III. Table V presents twenty words with the suffix *er* and Table VI presents twenty words with the suffix *ly*. The table is made up in exactly the same manner as Table III, is read in the same way and the results show the same thing.

From the data given above, the authors conclude, first, that in selecting words to make up the course of study in spelling each change in spelling should be incorporated as a separate word; second, that in teaching these words each separate form should be taught as a different word; third, that in making a scale for the measurement of spelling ability the spelling difficulty of each word must be discovered and the word assigned to its appropriate place in the scale.

TABLE VI

Difference between root and derived form of words adding ly. Grade II.

dear	2194	7	76	29	−47	
deep	16	9	53	23	−30	
glad	760	42	64	32	−32	−31
clear	75	10	44	13	−31	
fair	142	34	46	13	−33	−32
absolute		47	4	8	+ 4	
according	3	4	7	2	− 5	
actual	21	20	1	1	0	
annual	65	14	2	2	0	
awful	99	37	10	3	− 7	
certain	72	223	4	3	− 1	
complete	60	12	6	5	− 1	
correct	54	8	10	2	− 8	
direct	95	22	9	2	− 7	
entire	60	96	7	11	+ 4	
exact	22	24	0	4	+ 4	
extreme	8	13	5	0	− 5	
final	23	20	15	2	−13	
friend	485	12	16	19	+ 3	
general	140	16	5	0	− 5	− 2.5

COMPARATIVE TESTS OF HOME WORK AND SCHOOL WORK

W. H. HECK*

University of Virginia

The following tests were given on Tuesday and Wednesday, December 1 and 2, 1914, in the grammar and high school grades (5A–12) of the public school at Harrisonburg, Virginia. This town, with a population at that time of nearly 5,000 had long had a high average of comfortable homes and a general interest in public education. The students taking the test were representative of all classes in the community. The weather was good during the two days and the school work normal, the time having been selected so as to avoid inside or outside interference.

Grammar Grade Tests in Arithmetic

The test material was composed of four separate tests: (1 and 2) modifications of the very similar first and second forms of the Courtis test No. 7 in computation, and (3 and 4) the very similar first and third forms of the Courtis test No. 8 in reasoning problems. These tests represented practical arithmetic lessons, not artificial combinations, and the author had found them very satisfactory in his experiments on mental fatigue. The computation tests consisted of three examples in each of the four fundamental operations, and the reasoning tests consisted of eight problems using various combinations of the four fundamental operations. No fractions of any kind were involved, unless through a student's mistake. The time for any test need not have been as much as 15 minutes.

The total value of the examples in each of the computation tests was 16 (the individual values being ½, 1, or 2, with corresponding deduction for mistakes); the total value of the eight problems in each of the reasoning tests was eight (1 each); therefore, each student could have received a total of 48 points in the four tests (24 in his school work and 24 in his home work).

The half-year grades tested were 5A, 6B, 6A, 7B, 7A, and 8, the average age of the students by classes being approximately 12–15 years. These 6 grades were divided into two groups, as equally balanced as possible; the first group was composed of grades 8, 7B, and 6B, and the second group of grades 7A, 6A, and 5A. Only the class teachers were present with the children during the author's explanation, assignment and collection of each of the four test papers.

*Shortly after the completion of this article the author succumbed to an attack of pneumonia.—Ed.

There was no supervision of the student's work on the tests, and the teachers went on as usual with their regular instruction and discipline. Special supervision of this work by the class teacher would not have been typical of most study periods in school and would have given the school work undue advantage over the home work in the present experiment.

As rapidly as practicable after the opening of school on Tuesday, test 1 (computation 1) was assigned to the first group, and test 3 (reasoning 1) to the second group. The students were told that they could spend plenty of time in their study periods (when another division of the same class was reciting) to work out all the examples carefully; that the teachers would accept a reasonable excuse if the test hindered somewhat the preparation of the regular work; that the rating of the papers would not be done by anyone connected with the school and would not count against them individually; that the marks would be averaged by classes so as to show how well each class had profited by its arithmetic instruction; that any assistance of one student by another would be dishonest and would confuse the result. The students accepted the work gladly as a variation in the school routine and as a chance to show what they could do.

The papers were collected toward the close of the afternoon session, and test 3 (reasoning 1) was then assigned to the first group and test 1 (computation 1) to the second group, the work to be done by each student at home without help from anybody else and to be handed in after the opening of school next morning. Nothing at all was said about a comparison between home and school work, and the children did not show at any time during the series the slightest suspicion as to the final purpose of the tests. On Wednesday morning the six sets of papers were collected, and then test 4 (reasoning 2) was assigned to the first group and test 2 (computation 2) to the second group. These papers were collected toward the close of school, and test 2 (computation 2) was assigned to the first group and test 4 (reasoning 2) to the second group. These last papers were collected Thursday morning.

The first group did the first computation and the second reasoning test in school on Tuesday and Wednesday, respectively, while the second group was doing the first reasoning and the second computation test; and the first group did the first reasoning and the sec-

ond computation test at home on Tuesday and Wednesday, respectively, while the second group was doing the first computation and the second reasoning test.

The practice effect was thus so divided between the school work and the home work as to be neutralized and consequently in favor of neither. It was Schmidt's[1] failure to give due weight to the pratice effect in his methods and results, that furnished the initial suggestion to the author for the present experiments. The practice effect here, however, is slightly negative in the easy computation work (probably on account of increasing carelessness) and but slightly positive in the case of the more difficult reasoning problems. If we leave out of consideration for the moment or neutralize by a rearrangement of the marks, any possible difference between home and school work, we find that the average mark of all the students in the second computation test is 2.7 percent. less than that for the first, and the mark for the second reasoning test is 4.9 percent. more than that for the first.

Taking the 141 students or the six classes as a whole, the school work with its one computation and one reasoning test was as nearly equal as practicable to the home work with its one computation and one reasoning test, the only general difference being that between the school environment and the home environment. And a measurement of the influence of this difference in environment upon the work of school children was, of course, the main object of the present experiments.

The results were as follows:

TABLE I

Average Mark

Class	Students	School Work	Home Work	Difference
8	24	23.00	22.79	−0.21
7A	21	20.62	19.62	−1.00
7B	22	20.50	20.86	+0.36
6A	20	20.75	21.45	+0.70
6B	23	20.43	18.22	−2.21
5A	31	17.52	19.65	+2.13
Total	141	20.25	20.39	+0.14
Average of Class Averages	23.5	20.47	20.43	−0.04

[1]Archiv fur die gesammte Psychologie, 1904, vol. III, pp. 33–152; Zeitschrift fur Experimentelle Paedagogik, 1907, vol. IV., pp. 189–210, vol. V, pp. 23–50. These articles and the somewhat similar report by Meumann can be found in Heck's *International Source Book on Home Study*, to be published as a bulletin of the U. S. Bureau of Education.

It is easily seen from Table I that the difference is so close to 0 as to be moved a tiny fraction above or below by the method of calculation: 0.14 point, or seven-tenths of one per cent., in favor of home work in case of an average of all the students, and 0.04 point, or two-tenths of one per cent., in favor of school work in case of an average of class averages. Three of the grades (2 in the first group and 1 in the second) are slightly in favor of school work, and three (1 in the first group and 2 in the second group) are slightly in favor of home work—about as equal a balance as possible. It does seem that the two lowest classes, 6B and 5A, show preferences which are more than a negligible chance variation; but these perferences are almost equally balanced against each other.

Home work in arithmetic can be done as well as school work, if the children will to make it so; and objections on account of its supposed inferiority to school work do not seem to be justified. As a matter of daily routine the children may not know how or may not will to do either well. But it is precisely the function of good teaching, and of the home in co-operation, to meet such needs.

High School Tests in English Composition

On Tuesday morning the author assigned to the first high school group (consisting of 30 students from the first year and 11 from the fourth year class) a composition of between 500 and 1,000 words on the subject, "The Harrisonburg High School." The papers were to be collected just before the close of the afternoon session. The students were told to take as much time as they needed to do their best, although preparation for or even attendance at some class period might have to be sacrificed. Nothing was said as to making one or more copies, this being a matter for the student's own judgment and habit. It was explained that the marks on the papers were not to count for or against the students individually but were to show how well the different classes were profiting by their English instruction. These compositions were written during the usual study periods in the assembly hall and were not specially supervised by the principal in charge of the hall. After the collection of the papers a second composition, "The Town of Harrisonburg," was assigned to the first group, to be written independently and honestly by each student at home and handed in at the opening of school the next morning. In a similar way the first composition, "The Harrisonburg High School," was assigned as home work to

the second group (consisting of 21 students of the second year and 13 of the third year class). The two sets of papers were collected on Wednesday morning and the second composition was then assigned as school work to the second group. The few compositions which were handed in later, because they had to be finished or copied in school, were put together separately but without any comment or sign being made of the fact. Not only were these papers thrown out but so were the past or future test compositions by the same students, and also those by any student who wrote only one. On Thursday the students were surprised when told that the main purpose of the tests was a comparison between home work and school work.

It must be noted that the subjects for the two compositions were such that any student of high school grade ought to have had enough facts in mind to write a composition easily; certainly there was no need to consult outsiders or literature, a need that often brings injustice to some students who have inferior opportunities for such assistance. Not the getting of facts but the selection from the many at hand, and then the individual expression of ideas in regard to them—these were the real requirements and the two compositions were very nearly alike in their demands as to content and as to form.

The first composition was written in school by the first group and at home by the second group; the second composition was written at home by the first group and in school by the second group. This balancing was done primarily to neutralize the practice effect. However, the practice effect was really negligible in the present experiment, as the second composition was 0.2 per cent. less in form and 1.6 per cent. more in content than the first.

Although the grammar grade tests in arithmetic were easily marked and compared by the author, the case was far different and more difficult with the high school tests in English composition. In fact, these 150 papers were put aside on account of the pressure of other work until the last quarter of 1917, when Mr. W.C. Whitlock, editor of the students' magazine at the University of Virginia, was employed to mark the papers, so that a report on the tests could be included in the present report. After rejecting as futile the attempts to apply some so-called standard system of grading to these papers, the author decided on a double plan: one mark for

form—based on deductions for the mistakes in spelling, grammar, punctuation, handwriting, etc.; and one mark for content—the selection, organization, and interpretation of facts. (Both the formal and the material errors emphasized by Schmidt correspond to those included in the first mark here, as he neglected the more important and positive values of content.)

The marking for form was simple. Each mistake counted one and the total number was divided by the total number of words in the composition; then the resulting percentage was subtracted from 100, the remainder being the mark for form. Of course, this method has its deficiencies: some mistakes have little relation to words and the number of them; some words (the articles, for instance) are free from mistakes and are very frequently used; there may be more than one mistake in the same word; repetitions of the same mistakes are not of the same value as new mistakes; different mistakes have different practical values. Nevertheless, this method gives consideration to the necessity of proportioning mistakes to the length of the composition, as it would be manifestly unfair to compare *per se* the mistakes of form in one composition with those in another of twice the length. However, any class method that is exactly and impartially applied to all the compositions would serve the present purpose of comparing school and home work. As the average number of mistakes in each group is given in the table, other comparisons of these mistakes are made possible, if desired.

In regard to content, the best method of approach to exactness and impartiality in marking the papers seemed to be the reliance upon the judgment of one person as to the comparative merits of each and all of the papers. One trained mind reacting to 150 compositions on two clear and closely allied subjects gave more promise than any system planned to be used by many minds on many types and subjects of compositions. But the attainment of impartiality in this single judgment of content required several strict conditions, which were conscientiously carried out in the present case. (1) The mark for form was not to influence the mark for content, except in so far as the content was directly affected by the mistakes in form. (2) The marking was to be done by a person entirely ignorant of the grouping, order,etc., of the papers for comparison of home with school work. The papers were, therefore, shuffled into a confusing disarray before they were given out to be

marked. (3) The marking of each paper was to be done without any recollection of or reference to the previous content mark of the other paper by the same student. (4) The marking of each paper was to be done by a person fresh in mind, so that fatigue or boredom would not warp the judgment. Only a few papers were gone over at one time, and the marking of the 150 papers was extended over three months.

The results were as follows:

Class	Students	Number of Words		Number of Mistakes		Form Mark		Content Mark	
		School	Home	School	Home	School	Home	School	Home
Total	75	499	546	44	45	91.18	91.57	80.59	81.30
First Group	41	465	528	39	42	91.61	92.05	79.51	81.15
Second Group	34	541	567	51	49	90.57	91.36	81.88	81.62
Boys	31	476	550	51	53	89.29	90.36	78.81	80.74
Girls....	44	515	543	39	39	92.43	92.82	81.84	81.68
Reduced Total	64	492	534	45	45	90.85	91.57	80 53	81.16

Table II shows that the home compositions as a whole are but 0.4 per cent. better in form than the school compositions and but 0.9 per cent. better in content—differences so small as to be negligible in school practice. To be sure, the very small percentages in favor of the home compositions and the additional fact that these compositions are 9.4 percent. longer than the school compositions might imply slightly more freedom and more thoughtfulness of written expression in the home environment; but this suggestion has very little proof. (Schmidt believes that greater length of the compositions implies less thoughtfulness.)

Furthermore, the lack of any important difference between the home and the school compositions is as noticeable when the smaller groups are compared with each other. With a possible practice effect in their favor, the home compositions in the first group are 0.5 and 2.1 percent. better in form and content, respectively, than the school compositions; and with a possible practice effect the other way, the home compositions in the second group are 0.9 percent. better in form and 0.3 percent. worse in content. In

order to see whether the difference was greater or less in case of the boys as compared with the girls, averages were made for the 31 boys (15 in the first group and 16 in the second) and for the 44 girls (26 in the first group and 18 in the second). The home compositions of the boys were found to be 1.2 percent. and 2.4 percent. better in form and content, respectively, than their school compositions, and the home compositions of the girls were found to be 0.4 percent. better in form and 0.2 percent. worse in content than their school compositions. The smaller differences in the case of the girls are due to the proportionately better work of the girls in school. In order to make a total with the first and second groups equal in number of boys and of girls represented and in number of students in the classes balanced against each other, the first papers alphabetically in the various sets for boys and for girls were eliminated until there were 15 boys and 17 girls in each group and 21 first year and 11 fourth year students in the first group and 21 second year and 11 third year students in the second group. In case of this equality in the two groups, the 64 home compositions were only 0.8 percent. better in both form and content than the 64 school compositions. There seems no way of getting from the results of these tests any difference one way or the other between home and school compositions, which is large enough to affect school practice.

A significant fact is revealed by the amount of individual variation in the home as compared with that in the school compositions. The standard deviations from the average marks of the home compositions for form and for content are 3.29 and 4.06, respectively; while the standard deviations for the school compositions are 3.83 and 6.05, respectively. The deviations in the school compositions are 16.4 percent. greater in the marks for form and 49.0 percent. greater in the marks for content than are those in the home compositions. The S. D. for the home compositions by the boys are 2.94 and 4.14 for form and for content, respectively; while the S. D. for their school compositions are 3.86 and 7.34, respectively. The S. D. for the home compositions by the girls are 2.59 and 3.95 respectively, and for their school compositions 3.27 and 4.44, respectively. In every case there is greater deviation in the school compositions, and more especially so in the less mechanical phase—that of content. And the boys, with greater variation throughout than is the case with the girls, show even proportionately more in their school, as compared with their home, compositions

This result is further corroborated by comparisons between the standard deviations of the home and school work in arithmetic. The 6 classes (141 students) had a S. D. of 2.25 in their home computation test and of 2.36 in their corresponding school test; and an S. D. of 1.42 in their home reasoning test and of 1.81 in their corresponding school test. These deviations are necessarily smaller than in the case of the compositions, because the possible range is much smaller, a perfect paper in computation getting a mark of only 16 and a perfect paper in reasoning problems a mark of only 8. The difference between the deviations in the marks on the reasoning problems are especially important, as these tests required more mental work than did the computation tests.

This greater variation in the work of students in the more or less uniform environment of the school, at least externally, than in the many homes with their multitude of differences, is cause for surprise. Evidently the influence of the school's uniform environment is very far from being uniform, and the supposed objections to home work on account of the varied influences at home apply more forcibly to the school work. The school has been and probably will long continue to be in part an artificial environment, to which children make adjustments with widely varying degrees of completeness and naturalness, and to which the individuality of their responses is suppressed in the daily routine of mass instruction and discipline. But a test, with its demands for individual work, brings out not only the differences in the children *per se* but the exaggerated differences of their reaction to the school environment in which the work is done. On the contrary, in the shelter of the home the child is more at ease, his lifelong adjustment to the persons and to the things there have given him the chance to express himself more normally, even in work for the school. It might be said that the greater variations in the school environment correspond more closely to the natural variations among children and that at home there is a general tendency to restrict and reduce these variations; but this explanation seems in itself unnatural. As nearly all tests of the variability in children's work have been made in school, it is probable that the educationists are tending to exaggerate the amount of real difference among children.

Assignments for home work requiring equipment, literature, assistance, etc., which many homes of the students concerned cannot furnish, are manifestly unjust and should not be given in public

schools. But for sensible assignments, calling for independent work on topics previously discussed in school, the conditions in the majority of homes are more natural and, on the whole, about as favorable for individual work as the conditions in the majority of schools, however much improvement is needed in both environments. The life-long adjustments of the children have tended to neutralize, even to utilize the differences in homes in expressing the more or less general spirit of family life and in meeting therein the duties of home and even of school. Anyway, these differences are far greater to the outsiders looking in than to the insiders looking out. And school people are generally outsiders looking in.

As a final conclusion from the experiments in arithmetic and in English composition, chosen to represent two main and somewhat contrasted types of school requirements, we may say that home work and school work are or can be made about equal to each other, are about equally dependent on good teaching, and will therefore develop as the teaching develops. There is no use in pitting one element in this trinity against another, when the problems of clear instruction and assignment, creation of interest, and strict checking-up of results involve each one mutually.

COMMUNICATIONS AND DISCUSSIONS

PSYCHOLOGY IN THE UNGRADED ROOM OF THE VILLAGE SCHOOL

The need for careful, scientific school management is, in all probability, as great in village schools as elsewhere, yet nowhere is the difficulty of securing it so great. For example, the ratio of low normal and sub-normal children to the total population of the school is fairly constant everywhere. The village school of six or seven hundred pupils needs, then, the services of a psychologist and a capable teacher of sub-normal children quite as much as does the city school. But, under present conditions, the village school *cannot* have a psychologist; and if it has the capable teacher of sub-normal children, such a teacher must be developed from material at hand. When an ungraded room is established under such conditions, how may the superintendent and the board of education have any assurance at all that the services rendered are performed in a reasonably efficient manner?

The services rendered by the organization of the ungraded room may be summed up in a general way as follows: (1) selection, (2) classification, (3) instruction. It is not the purpose of this paper to explain any of these functions in the organization and instruction of an ungraded class, but to suggest one possible check for the selection and classification of the pupils in the ungraded class, namely, the coefficient of correlation between the mental ages of pupils as found by the application of a general intelligence test and the scores made by pupils on Trabue Language-Completion Scales.

During May and June of last year and September of this year, regular classroom teachers recommended for examination about sixty children. Fifteen of these were selected, eleven of whom had attained the mental age of eight years or more as shown by the Yerkes-Bridges Point Scale. These ages were later corrected so as to read as of October first and used in the correlation table given below.

During the first week of October, the Trabue Language-Completion Scales, B and C, were used in measuring pupils of grades 2, 3, 4, 5, 6, 7, 8. Medians for each grade and attained age were found and represented on a graph. The graph was used as a norm in determining the Trabue scale mental age from the score of an individual.

The eleven pupils in the ungraded room who had attained the mental age of eight years or more were measured by the Trabue Language-Completion Scales, B and C, administered on two different days two weeks apart. Two pairs of scores were identical, six showed a difference of one point, two a difference of two points, and one a difference of four points. The aggregate score on the first test was 89, while the aggregate score on the second test was

80. This is in accord with the facts as regards difficulty of the two scales. Where the score for a single individual differed, the average was taken as the true score, but it will be noticed that this reduced the scores from the first scores on the average less than half a point. The scores were then changed into mental ages by means of the norms described above. That these norms would, in general, be preferable to age norms taken from the St. Paul Survey, say, is certainly clear to most persons who will read this article. One specific reason is given. The measures were made in St. Paul in the middle of the school year of pupils who had attained the age at which they are reported in most cases by the first of the preceeding September. They would thus have a better opportunity to overcome the hard steps between sentences three and four, and four and five on the Trabue scales than would the children measured, in the main, soon after the attainment of the age at which they are reported. The following is the correlation table by the Pearson formula:

	Mental	Ages	Deviations					
Cases	Yerkes-Bridges	Trabue	x	y	xy	x^2	y^2	Excess of Yerkes-Bridges over Trabue age
1	8.4	7.7	−2.0	−1.7	3.40	4.00	2.89	.7
2	9.8	8.3	− .6	−1.1	.66	.36	1.21	1.5
3	11.4	10.9	1.0	1.5	1.50	1.00	2.25	.5
4	12.1	10.0	1.7	.6	1.02	2.89	.36	2.1
5	9.5	9.3	− .9	− .1	.09	.81	.01	.2
6	9.8	8.3	− .6	−1.1	.66	.36	1.21	1.5
7	8.9	7.9	−1.7	−1.5	2.55	2.89	2.25	.8
8	11.2	7.6	.8	−1.8	−1.44	.64	3.24	3.6
9	12.8	13.0	2.4	3.6	8.64	5.76	12.96	− .2
10	9.7	7.3	− .7	−2.1	1.47	.49	4.41	2.4
11	10.5	13.0	.1	3.6	.36	.01	12.96	−2.5
Av.	10.4	9.4	r	.66	0=1	.31	0=1.99	.96 Av.

The Trabue scales are simple to give, and the key makes it possible to score the results in an objective manner. A careful person of some scientific training can give the tests and score the papers just as accurately as a trained psychologist. Dr. Trabue reports in his monograph, *Language-Completion Scales*, that Dr. Scott of Northwestern University finds a correlation of .64 by the method of ranks

with a series of tests designed to measure the qualities thought to be necessary for success as efficiency engineers. Other workers since have more fully demonstrated the reliability of the Trabue scales as a measure of intelligence closely related to general intelligence. The rather high coefficient of correlation (r .66) reported here may then be taken (reasoning the other way around) to indicate that the selection and classification of the pupils in the ungraded room is fairly efficient. The author believes that other cautious superintendents may wish to resolve some of the doubts they now entertain as to the basis in mental testing of the ungraded rooms they have established, or are establishing, in a somewhat similar manner.

In partial explanation of the generally higher level of the mental ages found by the Point Scale, it might be added that an examination of the detailed records of pupils on the several tests shows that every case, except case eleven, tends to do a little less well where the test requires most language ability, and best where most thing-thinking is required. This tendency is extremely well marked in cases 4, 8, and 10, and re-enforces the familiar recommendation of hand work, and more hand work, for the pupil in the ungraded class.

Sup't of Schools, WM. L. CONNOR.
Republic, Michigan.

REPORT OF NEW CASES AND MORE RELIABE AGE NORMS OF INTELLIGENCE BY THE POINT SCALE FOR THE BLIND

Since the publication of *Mental Measurements of the Blind*, (Psychological Monographs, Number 89, April, 1916), a number of blind pupils in the state schools of Ohio and Kentucky have been rated by the Point Scale for the Blind.

Instead of the scores of only seventy-eight persons either totally blind or with such defects of vision that none could see the small (one inch) black wooden cubes used in weight discrimination when placed before him on white paper on the table at which he was sitting, I now have scores of one hundred and sixty such persons. Furthermore, the numbers of VII, VIII, and IX year-old children are greatly increased, a point of serious defect in Table II of the Monograph.

The Table of Scores, Table I, page 147, of the *Journal of Educational Psychology* for March, 1916, included scores of each pupil in the original mental survey of the Ohio State School for the Blind whose vision was too poor to enable him to be fairly tested by the Yerkes-Bridges Point Scale. Their visual acuity was not equal to

POINT SCALE SCORES. ONE HUNDRED AND SIXTY BLIND PERSONS WITH VISION TOO POOR TO SEE BLACK WEIGHT-DISCRIMINATION CUBES ON WHITE PAPER ON TABLE AT WHICH PERSON WAS SITTING. JULY 1, 1918

1	VI	VII	VIII	IX	X	XI	XII	XIII	XIV	XV	XVI	XVII	XVIII	XIX	XX	XXI	XXII	XXIII
	18	21	7	25	7	15	12	47	38	16	40	46	61	53	78	59	66	
		24	8	28	28	21	14	60	53	39	57	74	76	76	80	61	70	
		26	11	29	35	46	37	60	68	61	65	78	81	87	81	73	93	76
		28	12	34	36	47	44	64	74	62	69	83	87	90	84	83	97	81
		29	19	36	42	57	47	73	78	65	74	85	88	94	91	89	98	
		38	27	41	44	59	48	76	83	66	80	89	92		97	90		
		38	29	46	45	59	53	77	85	74	86	93	94		99	91		
		43	34	47	49	59	58	90	87	87	87		95			93		
		44	34	52	49	69	64		90	88	96							
		50	34	55	52	83	81		91	90								
			35	55	59	87	87		93	90								
			39		62		88											
			40		64													
			49		67													
			51		68													
			52		72													
			53		73													
2	18	33	34	41	49	59	50	68	83	66	74	83	87	87	84	86	93	
3	VI & VII 29		VIII & IX 34		X & XI 58		XII & XIII 60		XIV & XV 76		XVI & XVII 79		XVIII & XIX 87		XX & XXI 84			
4		34		46		57		76		71		85		85		89		
5		44	52	49	62	59	81	68.5	90	87	86.5	80.5	88	88.5	86	86.5	93	78.5

(1) Age nearest birthday when tested. (2) Medians 160 scores. (3) Medians by two year periods 160 scores. (4) Medians by two year periods 160 scores. (5) Medians by scores of 78 persons of Table II page 26 Monograph.

seeing the *pictures* well. But many of these persons could see the blocks quite well enough to perform the Knox Cube Test with no handicap. Such persons have different mental furniture from those who can merely see light and from the totally blind.

These new distributions of scores of the larger number of really blind pupils, which also reach down into the seventh year of life, with strength, give rise to a series of tentative norms which, from the numbers and distribution, are entitled to supplant all earlier presentations.

To the seventy-eight cases of Table II, page 26, *Psychological Monograph*, No. 89, the following additions have been made. Examiner, school, and date are given in each case.

Twenty-six	cases		T. H. Haines, Ohio State School for the Blind	Feb. 1916
Eleven	"	(11)	T. H. Haines, Ohio State School for the Blind	Dec. 1916
Thirteen	"	(13)	Miss A. C. Bowler, O. State Sch. for the Blind	Jan. 1917
Twenty-two	"	(22)	T. H. Haines, Ken. Inst. for the Blind	Feb. 1917
Fourteen	"	(14)	Miss A. C. Bowler, O. State Sch. for the Blind	Fall 1917

The same test material, examination sheets, and method of administration and scoring of tests have been used as described in the *Monograph*. Also the same method of distribution of scores by age.

The medians are given, as the most reliable tentative norms, especially for small numbers.

To show the improvement in conformity with the curve expected, if this point scale is a measure of intelligence, the medians, by years, of the seventy-eight cases of Table II of the Monograph are given in the last line of the table.

Both the series of medians by two-year periods show the expected advance during the years which intelligence tests from Binet onward have shown to be the years of marked development of that we call intelligence, which is measured by such tests.

The best norms in points, for the scale, suggested by this work to date, are as follows:

Years of Age	Expected Points
VI	18
VII	33
VIII	34
IX	41
X	49
XII	57
XIV	76

Ohio State University. THOMAS H. HAINES, M. D.

The Journal of Educational Psychology

INCLUDING EXPERIMENTAL PEDAGOGY, CHILD PHYSIOLOGY AND HYGIENE, AND EDUCATIONAL STATISTICS

Published monthly except July and August by
WARWICK & YORK, INC.
HERSHEY, PA. BALTIMORE, MD.

Board of Editors
WILLIAM CHANDLER BAGLEY, CARL E SEASHORE,
J. CARLETON BELL, Managing Editor, GUY MONTROSE WHIPPLE.

——COLLABORATORS——

Sydney Alrutz,	H. H. Goddard,	Will S. Monroe,	L. William Stern
J. R. Angell,	Joseph Jastrow	E. C. Moore,	G. M. Stratton,
J. W. Baird,	G. E. Johnson,	M. V. O'Shea,	G. D. Strayer,
Bird T. Baldwin,	E. A. Kirkpatrick,	W. B. Pillsbury,	E. L. Thorndike,
W. Van Dyke Bingham,	E. H. Lindley,	S. H. Rowe,	E. B. Tichener,
E. F. Buchner,	Anna J. McKeag,	Sante de Sanctis,	J. B. Watson,
W. H. Burnham,	D. P. MacMillan,	E. G. Sanford,	W. H. Winch,
Ed. Claparede,	Adolf Meyer,	W. S. Small,	H. K. Wolfe,
S. S. Colvin,	J. B. Miner,	E. D. Starbuck,	A. H. Yoder.
W. F. Dearborn,			

Manuscripts for publication, books or other material for review, and news items of interest to educational psychologists should be addressed to The Managing Editor, J. Carleton Bell, 1032A Sterling Place, Brooklyn, N. Y.

NOTICE TO SUBSCRIBERS. Notify the publishers, Warwick & York, Inc., of any change of address; as copies lost in transit because of subscriber's removal to another address will not be duplicated .

THE JOURNAL is issuing a series of Monograph Supplements for the publicaiion of more extensive researches in educational psychology. For details prospective contributors may apply to H. E. Buchholz, 10 East Center St., Baltimore, Md.

The price of the JOURNAL is $3.00 per year of ten monthly issues, payable in advance. Single issues are 50 cents each. Back volumes are $5.00 in original wrappers and $6.25 bound in library buckram. Foreign postage is 20 cents extra for Canada and 30 cents for other countries in the Postal Union.

Subscriptions and correspondence concerning business matters should be sent to Warwick & York, Inc.

VOLUME X MARCH, 1919 NUMBER 3

EDITORIAL

MENTAL TESTS AND COLLEGE ENTRANCE

The traditional college entrance examination has long been attacked as unsatisfactory by preparatory school teachers, by the candidates for entrance, and by college teachers. The theory of the examination was that a barrier was thereby set up to exclude from college those who were unprepared to do the work expected of them. Such expectation was capable of more precise definition when the bulk of college work consisted of Greek, Latin and mathematics, and when the preparatory school confined itself to the well established rudiments of these subjects. But with the introduction of electives into the college course and the later development of the group system, with the corresponding enrichment of the course of study in preparatory schools, and above all with the marvellous growth and increasing influence of the public high schools, it became more and more difficult to determine just what was to be expected of college freshmen and how their ability to meet these vague expectations could be revealed. The

preparatory teacher, especially the high school teacher, complained that the examination fettered him and prevented the kind of education which the child most needed. The high school graduates and their parents complained that the examination was arbitrary and capricious and did not allow the candidate to do himself justice. College teachers complained that under skillful coaching a very inferior student could be crammed sufficiently to pass the examination, and that classes were thus loaded with undesirable material. Thus any device that promised a broader and more reliable indication of the candidate's intellectual ability was likely to meet with a favorable reception.

Before the war tentative efforts had been made in a number of institutions to apply mental tests to freshman students, and to determine the correlation between the results of such tests and classroom work. The outcome of these efforts was not altogether satisfactory. Whether on account of the nature of the tests used or on account of the faulty technique in their administration, there was a distinct skepticism as to their reliability. The occurrences of the past two years, however, have given a tremendous impetus to applied psychology. The spectacular employment of mental tests on a large scale in the army has prepared the public mind for their use in other fields, and the necessity for selecting men for specific types of work has resulted in the development of tests that are said to have a high degree of reliability. For some time the Carnegie Institute of Technology has been using mental tests as an aid to the determination of the mental qualifications of freshmen, and now Columbia University announces that the traditional examinations will be supplemented by psychological tests to determine the candidate's ability to do college work. Professor Thorndike and his staff are at work on a battery of tests that will differentiate high school graduates into at least twenty grades or degrees of ability, and these tests will be available for other institutions after July 1.

All this is interesting and praiseworthy from the point of view of the scientific study of education, and it is to be hoped that the undertaking will receive the support that it deserves. At the same time the opinion may be hazarded that mere determinations of ability on the strength of general intelligence tests lack that precision and definiteness toward which scientific education is striving in diagnosis and prognosis. Such tests need to be supplemented by standard tests and scales in each of the high school subjects, so that the college authorities may have a complete diagnostic picture not only of the candidate's general intelligence (a vague term at best), but also of his proficiency in the work that he has attempted to do.

J. C. B.

NOTES AND NEWS

A correspondence course is announced by the University of California under the auspices of the California Children's Year Committee on the imagination, temperament, human cravings and play activities of children under twelve years of age. The material dealing with mental training has been prepared by Dr. Warner Brown, that on moral training by Professor C. E. Rugh.

Professor E. L. Thorndike, in his vice-presidential address before Section H, Anthropology and Psychology, American Association for the Advancement of Science, on "Scientific Personnel Work in the Army," recently published in Science, makes the following significant comments on the work of applied psychology: "Applied psychology is much more than cleverness and common sense using the facts and principles found in the standard texts. It is scientific work, research on problems of human nature complicated by conditions of the shop or school or army, restricted by time and labor cost, and directed by imperative needs. The secret of success in applied psychology or human engineering is to be rigorously scientific . . . . . When we took pains to compute the reliability coefficients of all our data before going further with them, we saved time in the long run. Every failure to check apparent meanings by objective correlations was disastrous. An unverified hypothesis may possibly be a relatively harmless luxury if all one does with it is to think; to act on it is a grave danger . . . . . It is relatively easy to be scientific when you can direct your talent in any one of ten thousand directions; yourself asking the questions for which you proceed to find answers! Psychology applied to the complicated problems of personnel work represents scientific research of the most subtle, involved, and laborious type."

Dr. Otis W. Caldwell, of the Lincoln School, New York City, and Mr. Stuart A. Courtis, of Detroit, have discovered a survey of the schools of Boston, Mass., made in 1845, and have printed an edition of 50,000 copies of questions in American history, geography, arithmetic and other subjects from the tests used in the survey. They desire to enlist the co-operation of the schoolmen of the country in giving these tests to eighth grade pupils some time during the last month of the present school year. The cost for a class of 30 pupils including record sheets and postage both ways will be $.65. The time required for giving the tests will be about 100 minutes of class time. Those who are willing to participate in this investigation are asked to write at once to S. A. Courtis, 51 Shelby Street, Detroit, Michigan. A certain number of the tests will be distributed free of charge.

At the February meeting of the New York Society for the Experimental Study of Education Mr. S. A. Courtis, director of the bureau of educational research, Detroit, Michigan, gave an illustrated lecture on "The Development of Standard Tests in Education, and Their Application to Problems of Classification and Diagnosis."

The new Prussian socialist Kultus-Minister has already introduced sweeping changes in the German school system. Among the 32 points of his program, as published by the *London Times*, there is to be complete separation of religion and religious authorities from the schools, all class schools are to be abolished and the "uniform" school maintained, co-education is more generally introduced, the teaching of history is re-organized, discussions of social policy are encouraged instead of forbidden, examinations are to be reduced, and teachers are to have representation in the government and in school administration.

The New York Branch of the American Psychological Association held a meeting on February 24, at which the following papers were read: "Correlations between the Standard Revision of the Binet Test and Performance Tests," by Dr. L. I. Stecher; "Examination for Emotional Instability," by Professor R. S. Woodworth; "The Mental State of Hysterics," by Professor H. L. Hollingworth.

Dr. Charles Barr Robertson, director of the division of university extension in the University of Pittsburgh, and one of the most popular school men in the State of Pennsylvania, died on February 4, after a month's illness.

Professor Alexander Inglis, of Harvard University, has been appointed by the Virginia State Board of Education to make an educational survey throughout the state.

Professor Irving King, of the University of Iowa, is now supervisor and organizer of the army convalescent school at Camp Travis, Texas. The Camp Travis Khaki College, a Y. M. C. A. school for soldiers, was organized during Dr. King's service in the southern cantonment.

Captain Garry C. Myers, Sanitary Corps, U. S. Army, is chief of educational service and morale officer at General Hospital No. 31, the former Indian School of Carlisle, Penna. Captain Myers is on leave of absence from the Brooklyn Training School for Teachers, New York City.

Captain R. H. Wheeler, professor of psychology in the University of Oregon, who has been conducting psychological tests in the army, has returned to take up his work at the university.

PUBLICATIONS RECEIVED

CARTER ALEXANDER. *School Statistics and Publicity.* Boston: Silver, Burdett and Company, 1919. Pp. xix, 332.

This book is intended for the school superintendent. It is designed to assist him in presenting the needs of his school system in the most effective and convincing manner. "The school superintendent must be a publicist. He must make reports to the public. In many places, for the next decade at least, he must fight as hard as any officer in the trenches to ward off the incessant and fierce attacks made upon his school appropriations by politicians and hardpressed but unthinking tax-payers. For warding off or beating back such attacks, his most effective weapons will be reports containing simple but skillful statistical devices for presenting the czaims of the school children." The book makes valuable suggestions for the collection of statistical data; indicates the significance of scales and tables of distribution; gives an admirably lucid discussion of central tendency, of deviation, and of correlation; and devotes almost one hundred pages to the discussion of devices for presenting statistics to the public in the most striking and attractive form. The wide-awake superintendent will find it extremely helpful and suggestive.

CHARLES R. ALLEN. *The Instructor. The Man and the Job.* Philadelphia: J. B. Lippincott Company, 1919. Pp. vii, 373. $1.50.

The instructor here referred to is the industrial instructor, whose business it is to take the crude, untrained, raw labor recruit in the factory and gradually develop such skill and efficiency as the man seems capable of attaining. This task has usually been left to the foreman, who has acted as though it were best accomplished by bullying and physical force. But managers and employers are coming to realize that this is a very wasteful method of procedure, and are turning to the specialist in handling men for more efficient methods. The author was formerly assistant superintendent of training in the United States Shipping Board, and has here set forth the results of his observations and experience.

WILLIAM H. ALLEN AND CLARE KLEISER. *Stories of Americans in the World War.* New York: Institute for Public Service, 1918. Pp. 176.

The authors have gathered together interesting personal accounts, newspaper reports, official citations, poems, and other material to give a composite picture of the part played by America in the war. It is excellent supplementary reading for the upper elementary grades.

WILLIAM H. ALLEN AND CARROLL G. PEARSE. *Self-Surveys by Teacher-Training Schools.* Yonkers-on-Hudson: World Book Company, 1917. Pp. xvi, 207. $2.25.

Critical self-analysis has not been a marked characteristic of American education. Especially the important function of teacher-training has been allowed to drift on the uncertain waves of tradition and local caprice. The present book is superficial and impressionistic, but it gives many valuable hints to live training school principals, and points the way to a more thorough and scientific study of the training of teachers.

J. MACE ANDRESS. *Health Education in Rural Schools.* Boston: Houghton Mifflin Company, 1919. Pp. xii, 321. $1.60.

The poet and the back-to-the-farm enthusiast may laud the superior health-giving influences of country life as compared with city life, but every careful study that has been made shows the lamentable neglect of the simplest rules of hygienic living in our rural populations. The physical examinations for army service showed that the country boy was just as defective as the city boy, and country school children have little attention paid to their health. The present volume is designed to adapt accepted hygienic teachings and practices to the needs of the rural school teacher. It is written in a simple, direct style, yet is thoroughly scientific and up-to-date, and contains valuable lists of supplementary books. The final chapter presents an interesting point scale for the measurement of the teacher's work in hygiene, and the application of the scale to typical teachers one of whom gets a score of 16 points, another a score of 84 points, and the third a score of 133 points. The general use of this scale would give us some interesting data on health teaching in the schools.

SHOLAM ASH. *The God of Vengeance, A Drama in Three Acts.* Boston: The Stratford Company, 1918. Pp. xiv, 99. $1.00.

A translation of a Yiddish drama portraying the struggle of ideals in the heart of even the most debased. The scene is laid in a brothel in the Jewish quarter of a Russian town.

FRANK P. BACHMAN AND RALPH BOWMAN. *The Gary Public Schools. Costs.* New York: General Education Board, 1918. Pp. 86, with thirty infolded tables. Twenty-five cents.

One of the most thorough and elaborate studies of school costs that we have available. The Gary school shops have been widely claimed to be self-supporting. It was soon found that such a claim could rest only on a very confused and irregular system of accounting. The surveyors were obliged to go back to the original vouchers and to evaluate the output of the shops for several years. As a result they conclude that if all the shops be taken together they are only 54 percent. self-supporting, while if only the essentially productive shops are considered by themselves there are 69 percent. self-supporting.

FRANKLIN BOBBIT. *The Curriculum.* Boston: Houghton Mifflin Company, 1918. Pp. viii, 295. $1.50.

Social,commercial, industrial and political life are changing very rapidly, but our educational program lacks thorough-going adjustment to these changes. What should constitute the course of study for the different grades from the kindergarten to the college, and what should be the relative proportions of the subjects included? "For a long time we have been developing the theory of educational method, both general and special, and we have required both teachers and supervisors to be thoroughly cognizant of it. Recently, however, we have discerned that there is a theory of curriculum formation that is no less extensive and involved than that of method, and that it is just as much needed by teachers and supervisors." The present volume offers an introductory survey of the problems involved in curriculum making. There is a preliminary discussion of the ends to be served by the curriculum, a study of the technical training necessary for occupational efficiency, an analysis of education for citizenship, an examination of the bases of education for physical efficiency, an inquiry into education for leisure, and a consideration of the linguistic demands of social intercourse. One of the most interesting and stimulating chapters in the book is that on reading as a leisure occupation.

EMORY S. BOGARDUS. *Essentials of Social Psychology.* Los Angeles: University of Southern California Press, 1918. Pp. 159.

This is a kind of outline or syllabus of problems in social psychology as developed in the author's classes. There are chapters on the psychological basis of social studies, the instincts, the characteristics and operation of imitation, temporary and permanent groups and their influence on behavior, and the psychology of invention, leadership and social control. Each chapter is followed by an extended list of problem-questions and a selected bibliography.

SANGER BROWN, II. *Sex Worship and Symbolism of Primitive Races.* Boston: Richard G. Badger, Pp. 145. $3.00.

This simple presentation of a fundamental motive should appeal to everyone who is interested in mental evolution. The historical portion of the book gives a description of sex worship which had its origin in the mind of primitive man, but which continued its influence (unrecognized for the most part) through the past ages down to the present day. A parallel is drawn between the history of this motive in the collective mind of the race and the influence of the sex motive in the life of the normal individual.

ANTON CHEKHOV. *Nine Humorous Tales.* Boston: The Stratford Company, 1918. Pp. ix, 60. Twenty-five cents.

These delightful tales are a relief from the sinister, dark, morose character of most modern Russian fiction. They remind one strongly of O. Henry in the directness of their style and the whimsical nature of their development.

HUBERT GUY CHILDS. *An Investigation of Certain Phases of the Reorganization Movement in the Grammar Grades of Indiana Public Schools.* Bloomington, Ind.: Indiana University Book Store, 1918. Pp. viii, 187. $1.50.

This is an investigation of the Junior high school, based upon a study of thirty-five so-called junior high schools, thirty-five typical departmental schools, and twenty-three non-departmental schools. Comparisons were made as to standards, programs of study, provision for individual differences, revised methods, social organization, costs, achievements in specific subjects, and retention of pupils in school. In many respects the junior high schools show improvement over the elementary schools in flexibility of promotion, work in practical arts, and wider opportunities for electives. The Woody Arithmetic Tests, Series B, Multiplication, the Ayres Spelling Scale, the Thorndike Reading Scale, Alpha 2, and the Thorndike Visual Vocabulary Scale were given to the pupils of schools in two counties which differed widely in the time given to English and arithmetic. Except for the spelling, the schools giving little time to these subjects made practically as good a showing as those that gave twice as much time to them.

SHERWIN CODY. *Commercial Tests and How to Use Them.* Yonkers-on-Hudson: World Book Company, 1919. Pp. vii, 216. Ninety-nine cents.

This little volume gives an account of the origin and use of the National Business Ability Tests. The chapter on "Principles of Scientific Tests" has a rather full account of the Alpha Army Tests and the appendix contains a comparison of the performance of the same pupils with arithmetic tests of the National series and the Courtis Tests, Series B. There are detailed directions for giving and scoring the tests, and the results of their application in various communities.

MAXIM GORKI. *Stories of the Steppe.* Boston: The Stanford Company, 1918. Pp. vii, 59. Twenty-five cents.

This is a volume of "The Universal Library" in which the publishers present the best literature of all nations in neatly bound volumes at twenty-five cents each. The series already contains eight or ten numbers. The present volume contains three of Gorki's vivid and poignant short stories.

SAMUEL HAMILTON. *Hamilton's Standard Arithmetics.* Cincinnati: The American Book Company, 1918. Book I, Pp. 256+xvi; Book II, Pp. 300+xx: Book III, Pp. 374+xxii.

These texts are generously illustrated, are well supplied with devices for drill and short cuts, and offer an abundance of practical problems.

ALICE M. KRACKOWIZER. *Projects in the Primary Grades.* Philadelphia: J. B. Lippincott Company, 1919. Pp. ix, 221. $1.28.

The traditional elementary school is admittedly formal and artificial. Yet abundant materials for real education lie all about children waiting the proper organization. In this book the author makes happy suggestions for utilizing the child's play activities, his social experiences, and his interest in nature, in building up preliminary habits for the more formal subjects of the school. Teachers of small children will find the book very helpful.

PAUL J. KRUSE. *The Overlapping of Attainments in Certain Sixth, Seventh and Eighth Grades.* New York: Teachers College, Contributions to Education, No. 92, 1918. Pp. 91.

This is an intensive study of the behavior of a limited number of pupils to a large number of mental and educational tests. The 860 pupils were distributed fairly evenly among the three grades, and the tests used were the Woody Arithmetic Tests, three arithmetical problem tests, four Trabue Completion tests, two composition tests, Thorndike Reading Scale Alpha 2, a spelling test, four opposites tests, seven other association tests, three directions tests, and two visual vocabulary tests. The author finds great overlapping from the results of single tests, but this overlapping is reduced materially in the composites of a number of tests of the same trait, and still further reduced in composites combining the measures of different traits. The coefficients of reliability of the arithmetic and English composites were very high, approximately .85.

ARTHUR W. KALLOM. *Arithmetic. Determining the Achievement of Pupils in Common Fractions.* Boston: Department of Educational Investigation and Measurement, Bulletin No. 15, 1918. Pp. 38.

This bulletin reports the results of investigations extending the previous study of addition of fractions to multiplication and division. An analysis was made of the types of mental processes involved in multiplication and division of fractions, and tests were constructed to reveal proficiency in these processes. The tests were given to 3513 pupils of grades VI–VIII in 95 classes. The results are tabulated to show the medians for speed and accuracy in each test for each grade, the percentage of failures in each test, and the sources of error in the attitude of the pupils. The bulletin closes with a valuable plan of diagnosis to be used by the teacher.

RICHARD SWANN LULL, EDITOR. *The Evolution of the Earth and its Inhabitants.* New Haven: Yale University Press, 1918. Pp. xi, 208. $2.50.

This is a series of lectures delivered before the Yale Chapter of the Sigma Xi by specialists in various fields of evolution. It includes "The Origin of the Earth"

by Joseph Barrell; "The Earth's Changing Surface and Climate," by Charles Schuchert; "The Origin of Life," by Lorande Loss Woodruff; "The Pulse of Life," by Richard Swann Lull; and "Climate and the Evolution of Civilization," by Ellsworth Huntington. The book presents an excellent survey of the most recent findings of various sciences on the subject of evolution.

INEZ N. MCFEE. *The Teacher, the School and the Community.* New York: The American Book Company, 1918. Pp. 256. $1.24.

This book is intended as a "guide, philosopher and friend" for the new teacher in the ungraded rural school. For the graduate of one of our better normal schools or training schools it would be entirely superfluous, but unfortunately there are thousands of young teachers all over this country who have had so little preparatory training that the book would be a godsend to them. There are numerous references to teachers' helps that will prove valuable to those seeking such material.

ALBERT E. MCKINLEY, CHARLES A. COULOMB, AND ARMAND J. GERSON. *A School History of the Great War.* Cincinnati: The American Book Company, 1918. Pp. 192.

Already we are far enough away from the great tragedy to make a school history of it very welcome. The present study is an admirable story of the war and will make a strong appeal to young people. The narrative is brought down to the signing of the armistice on November 11, 1918.

BRANDER MATTHEWS. *An Introduction to the Study of American Literature.* Cincinnati: The American Book Company, 1918. Pp. 268.

A new and revised edition of this valuable little manual. The author still holds to the principle of excluding all living writers, thereby securing more space for the consideration of the earlier literature, but robbing his readers of the benefit of his opinions on contemporary movements.

EDWARD J. MENGE. *The Beginnings of Science, Biologically and Psychologically Considered.* Boston: R. G. Badger, 1918. Pp. 256. $2.00.

This series of papers presents a Roman Catholic reaction to the trends of modern science. There are treatises on biology, psychology, genetics, metaphysics, logic, evolution, vitalism and the ideal. The author's interpretations are scarcely in accord with the views of most leaders in scientific thinking.

WALTER SCOTT MONROE. *Measuring the Results of Teaching.* Boston: Houghton Mifflin Company, 1918. Pp. xviii, 297. $1.60.

This is perhaps the best presentation of the results of educational measurements to date. The book embodies much of the material in the author's previous *Educational Tests and Measurements*, written in collaboration with Mr. James C. DeVoss and Dr. Frederick J. Kelley, but omits many of the tests mentioned in that book, and adds some others. The omissions enable the author to give a more extended treatment to the problems of diagnosis and to the correction of the defects revealed by the tests. On the other hand the emphasis on a few types of tests leaves the book unsatisfactory as a comprehensive account of the present status of educational measurements. In the very beginning, however, the author disavows any intention of producing a fundamental treatise on the subject, but has aimed rather at "a text which will help teachers to use standardized tests to the greatest advantage." In the chapter on reading only the Monroe Silent Reading tests, the Courtis Silent

Reading test, the Thorndike Visual Vocabulary, and the Gray Oral Reading test are discussed, but over 50 pages are devoted to the correction of defects in reading. The appendix contains valuable suggestions for procuring the most important standard tests.

CHARLES CLINTON PETERS. *Human Conduct. New York:* The Macmillan Company, 1918. Pp. xii, 430. $1.30.

The sub-title of this book is *A Textbook in General Philosophy and Applied Psychology for Students in High Schools, Academies, Junior Colleges, and for the General Reader.* Many educators have been asking, "Why should high school science be limited to physical science? Why would not mental science offer valuable educative material to the adolescent?" There is no essential reason, other than tradition, for the restriction of psychology and sociology to college students. Already progressive schools are introducing courses in economics. The present book selects those phases of psychology and logic which would make the strongest appeal to young people and sets them forth in direct and simple language. The first part of the book is concerned chiefly with illusions, and applies the psychology of illusions to an explanation of the way in which misunderstandings arise. Among other topics discussed are apperception and its significance for the individual and the race; the solution of problems and clear thinking; the search for the causes of things; the control of conduct through ideas; the role of memory and imagination; attention and habit formation; and the selection of one's life work. While the book is somewhat diffuse and discursive, the reading of it would be of distinct benefit to any high school pupil.

JOHN C. STONE. *The Teaching of Arithmetic.* Chicago: Benjamin H. Sanborn and Company, 1918. Pp. v, 262. $1.32.

This is a very clear and simple presentation of the essential facts of arithmetic. The first chapters deal with the fundamental number combinations, chapter eight suggests many games that may be utilized in drill work, and subsequent chapters discuss common and decimal fractions, percentage, denominate numbers, analysis and solution of problems, and planning the work in arithmetic. The final chapter is entitled "Measuring Results," and contains a summary of the work of Rice, Stone, Courtis, Woody, and the tests used in the Cleveland Survey.

C. W. TABER. *The Business of the Household.* Philadelphia: The J. B. Lippincott Company, 1918. Pp. xii, 438. $2.00.

This volume is in keeping with the high standard of excellence set by previous volumes in the Lippincott's Manuals. Part I deals with fundamental principles of household finance, including the family income, bank and household accounts and the family budget. Part II discusses the necessities of the budget, such as rent, fuel, light, taxes, insurance, and marketing. Part III considers such problems as service in the home, insurance, savings, investments, old age, and the cultural wants of the family. Part IV takes account of the legal and business status of the family, treating of real estate titles, inheritances and wills, laws affecting the family, and business principles in the home. The discussions are thorough and scholarly, and based upon the best authorities.

ORDWAY TEAD. *Instincts in Industry.* Boston: Houghton Mifflin Company, 1918. Pp. xvii, 222. $1.40.

This is not a psychological study, but an attempt to analyze and present certain factors of the labor situation on the basis of a scheme of instincts. The author

selects what he considers ten basic instincts, and under each rubric calls attention to certain crucial problems in industrial life. What he has to say about industrial problems is interesting and pertinent, but for the psychologist who is at all critical in the use of his terms the attempt to explain anything by the employment of such terms as the instinct of workmanship, the instinct of self-assertion, the instinct of submissiveness, and the instinct of the herd is pure word juggling. The book is another illustration of the pernicious effect upon economic writers of the craze for instincts started by James and fostered by McDougall and other social psychologists.

FREDERICK J. TEGGERT. *The Processes of History.* New Haven: Yale University Press, 1918. Pp. ix, 162. $1.25.

This is an attempt to do for human history what biologists are engaged in doing for the history of the forms of life. The great problem of history is to determine how man everywhere has come to be as he is. The various general theories that have been proposed are carefully examined, the geographical factor in history is given due consideration, and the age-long conflict between individualism and collectivism is shown to be the fundamental human element in history. The book is particularly valuable for its point of view and for the method of attack proposed for the problems of history.

WILLIS MASON WEST. *History of the American People.* New York: Allyn and Bacon, 1918. Pp. xxvi, 729+44. $1.75.

This text, intended for use in high schools, emphasizes the development of industrial and social life in America. According to the author the following points have received special attention: (1) The historical grounds for friendship between America and England, in spite of old sins and misunderstandings. (2) The meaning of the West in American history. (3) The heroic labor movement of 1825–1840, usually ignored. (4) The long conflict between intrenched "privilege" and the "progressive" forces in State and Nation. A final chapter on the World War brings the narrative down to January, 1918. The book is profusely illustrated with pictures, maps and diagrams, and the liberal use of paragraph numbers and headings and bold face type brings out the statements which the author considers of prime importance.

JOHN F. WOODHULL. *The Teaching of Science.* New York: The Macmillan Company, 1918. Pp. xiv, 249. $1.25.

This collection of papers and addresses touches upon a large number of topics connected with the teaching of science, and covers a period of over twenty years. Among the subjects discussed are the enrichment of high school courses in physics and chemistry to give them greater cultural content and significance, the evils of specialization in high school science, the effects of college entrance requirements in natural science, learning from experience, and the project method in the teaching of science. In some measure the book serves as a history of the natural science movement in recent education.